A.G. KUROSCH

CURSO DE ALGEBRA SUPERIOR



EDITORIAL MIR

А. Г. КУРОШ

КУРС ВЫСШЕЙ АЛГЕБРЫ

ИЗДАТЕЛЬСТВО «НАУКА»

На испанском языке

A. G. KUROSCH

CURSO de ALGEBRA SUPERIOR

Traducido del ruso por

EMILIANO APARICIO BERNARDO,

Candidado a Doctor en Ciencias Físico-Matemáticas,

Catedrático de Matemáticas Superiores.

EDITORIAL MIR • MOSCU 1968

CDU 512.8 (075.8) = 60

INDICE

PALABRAS DE PRESENTACION

La versión castellana de la obra del profesor A. G. Kurosch «Curso de álgebra superior» que ofrecemos al lector, es el primer libro del autor que se traduce al español.

El conocimiento del álgebra superior es indispensable para la formación matemática del estudiante que ha decidido consagrarse al estudio de las matemáticas. El presente libro marca un camino relativamente corto para pasar del álgebra elemental al estudio de los métodos abstractos del álgebra moderna.

En los primeros capítulos se estudian detalladamente los determinantes y sistemas de ecuaciones lineales, se introducen los números complejos y las operaciones sobre las matrices, y se hace una exposición de la teoría de los polinomios y formas cuadráticas. En los capítulos VII y VIII, el autor nos da una idea primordial del álgebra lineal. En el capítulo X vemos que el álgebra lineal, el álgebra de los polinomios y las funciones racionales pueden generalizarse para el caso de un campo fundamental arbitrario. Precisamente en este capítulo, el autor nos enseña los principios del álgebra moderna. Aquí nos encontramos con los conceptos importantes de anillo y campo. Estos conceptos permiten exponer con mayor generalidad la teoría de los polinomios en varias indeterminadas, suponiendo que los coeficientes de estos polinomios pertenecen a un campo fundamental arbitrario. A continuación, las matrices polinomiales también se estudian sobre un campo fundamental arbitrario y se aplican para la elaboración de la teoría de las matrices de Jordan. El último capítulo está dedicado a los grupos; éste es el comienzo de una rama muy importante del álgebra moderna, denominada teoría de los grupos.

El autor de este libro es un gran especialista en teoría de grupos. Su libro «Teoría de los grupos», desempeñó un papel muy importante en el desarrollo de las investigaciones sobre este tema en la Unión Soviética. Hace unos años, el prof. A. G. Kurosch publicó una original obra, titulada «Lecciones de álgebra general», que fue favorablemente acogida por los algebristas soviéticos.

El prof. A. G. Kurosch es jefe de la cátedra de álgebra superior de la Universidad de Moscú desde el año 1949.

El presente libro es un compendio de álgebra superior que comprende los conocimientos de esta ciencia obligatorios para los estudiantes de matemáticas de la Universidad de Moscú. Desde la aparición de su primera edición en ruso, en el año 1946, ya ha sido reeditado ocho veces. En la Unión Soviética éste es uno de los mejores libros sobre el tema considerado. Esperamos que tenga buena acogida en los países de habla hispánica.

Agradeceremos al lector sus observaciones sobre la presente traducción, que trataremos de tener en cuenta en el futuro.

Moscú. Febrero de 1968. *E. Aparicio Bernardo*

CAPITULO I

SISTEMAS DE ECUACIONES LINEALES. DETERMINANTES

§ 1. Método de eliminación consecutiva de las incógnitas

Comenzamos el curso de álgebra superior con el estudio de los sistemas de ecuaciones de primer grado con varias incógnitas o, como suele decirse, *de los sistemas de ecuaciones lineales* *.

La teoría de los sistemas de ecuaciones lineales origina una amplia e importante rama del álgebra, el álgebra lineal, a la que están dedicados una gran parte de los capítulos de este libro y, en particular, los tres primeros. Se supone que son reales los coeficientes de las ecuaciones que se consideran en estos tres capítulos, los valores de las incógnitas y, en general, todos los números que aparecen. En realidad, todo el contenido de estos capítulos se generaliza, palabra por palabra, al caso de números complejos arbitrarios, ya conocidos por el lector en el curso de la escuela media.

A diferencia del álgebra elemental, aquí se estudian los sistemas con un número arbitrario de ecuaciones e incógnitas. Además, suponemos que el número de ecuaciones del sistema *no* coincide con el número de incógnitas.

Sea dado un sistema de s ecuaciones lineales con n incógnitas. Convengamos en emplear las siguientes notaciones: las incógnitas las designaremos con la letra x con subíndices $1, 2, \ldots n$: $x_1, x_2, \ldots, x_n$; supondremos que las ecuaciones están numeradas así: la primera, la segunda, . . ., la s-ésima; el coeficiente de la incógnita x_j en la i-ésima ecuación, se señalará mediante a_{ij}**; finalmente, el término independiente de la i-ésima ecuación se designará con b_i.

---

* Esta denominación se debe a que, en la geometría analítica, una ecuación de primer grado con dos incógnitas determina una recta en el plano.

** Por consiguiente, se emplearán dos subíndices, el primero de los cuales indicará el número de la ecuación, y el segundo, el número de la incógnita. Para abreviar, estos índices no se separarán con una coma; claro que, en el caso de a_{11}, no se debe leer «a once», sino «a uno uno», y en el caso de a_{34}, no se debe leer «a treinta y cuatro», sino «a tres cuatro».

Nuestro sistema se escribirá ahora en la forma general siguiente:

$$\left.\begin{array}{l} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n = b_1, \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n = b_2, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ a_{s1}x_1 + a_{s2}x_2 + \ldots + a_{sn}x_n = b_s. \end{array}\right\} \quad (1)$$

Los coeficientes se pueden colocar formando un cuadro

$$\begin{pmatrix} a_{11}a_{12} & \ldots & a_{1n} \\ a_{21}a_{22} & \ldots & a_{2n} \\ \ldots & \ldots & \ldots \\ a_{s1}a_{s2} & \ldots & a_{sn} \end{pmatrix}, \quad (2)$$

denominado *matriz de s filas* y *n columnas*; los números a_{ij} se llaman *elementos* de la matriz*. Si $s = n$ (o sea, que el número de filas es igual al número de columnas), se dice que la matriz es *cuadrada* y *de orden n*. La diagonal de esta matriz que une el ángulo superior izquierdo con el ángulo inferior derecho (o sea, formada por los elementos a_{11}, a_{22}, ..., a_{nn}), se llama *diagonal principal*. Una matriz cuadrada de orden n se llamará *matriz unidad de orden n*, si todos los elementos de su diagonal principal son iguales a la unidad, y todos los elementos que están fuera de esta diagonal son iguales a cero.

Se llama *solución* de un sistema de ecuaciones lineales (1) a un sistema de n números $k_1, k_2, \ldots, k_n$, en el que cada una de las ecuaciones del sistema (1) se convierte en una identidad, después de haber sustituido en ella las incógnitas x_i por los números correspondientes k_i, $i = 1, 2, \ldots, n$**.

Un sistema de ecuaciones lineales puede no tener solución alguna, y entonces se llama *incompatible*. Tal es, por ejemplo, el sistema

$$x_1 + 5x_2 = 1,$$
$$x_1 + 5x_2 = 7;$$

los primeros miembros de estas ecuaciones son iguales, mientras que los segundos son distintos. Por lo tanto, ningún sistema de valores de las incógnitas puede satisfacer simultáneamente a las dos ecuaciones.

Si el sistema de ecuaciones lineales tiene solución, se llama *compatible*. Se dice que un sistema compatible es *determinado*, si posee una solución única (en el álgebra elemental solamente se estudian

* De este modo, si la matriz (2) se examina sin relación con el sistema (1), el primer subíndice del elemento a_{ij} indica el número de su fila, y el segundo, el número de su columna.

** Hay que subrayar, que los número $k_1, k_2, \ldots, k_n$ forman *una* solución del sistema y no n soluciones.

tales sistemas), e *indeterminado*, si tiene más de una solución. En este caso, como veremos más adelante, hay una infinidad de soluciones. Así, el sistema

$$\left.\begin{array}{r} x_1+2x_2=7, \\ x_1+x_2=4 \end{array}\right\}$$

es determinado y $x_1 = 1$, $x_2 = 3$ es una solución. Por el método de eliminación de la incógnita, se puede comprobar fácilmente que esta solución es única. Por otra parte, el sistema

$$\left.\begin{array}{r} 3x_1-x_2=1, \\ 6x_1-2x_2=2 \end{array}\right\}$$

es indeterminado, puesto que tiene infinitas soluciones de la forma

$$x_1=k, \quad x_2=3k-1, \tag{3}$$

donde el número k es arbitrario. Con las soluciones obtenidas por las fórmulas (3) se agotan todas las soluciones de nuestro sistema.

El problema de la teoría de los sistemas de ecuaciones lineales consiste en la elaboración de métodos que permitan establecer si es compatible o no un sistema dado de ecuaciones, y en caso de compatibilidad, indicar el número de soluciones y señalar un método para hallar todas ellas.

Comenzaremos por el método más cómodo para hallar prácticamente las soluciones de los sistemas con coeficientes numéricos, es decir, *con el método de eliminación consecutiva de las incógnitas o método de Gauss**.

Hagamos primero una observación. A continuación, tendremos que hacer las siguientes transformaciones del sistema de ecuaciones lineales: ambos miembros de una de las ecuaciones del sistema, multiplicados previamente por un mismo número, se van a restar de los miembros correspondientes de otra de las ecuaciones del sistema. Supongamos, por ejemplo, que ambos miembros de la primera ecuación del sistema (1), multiplicados por el número c, se restan de los correspondientes miembros de la segunda ecuación. Obtendremos un nuevo sistema de ecuaciones lineales:

$$\left.\begin{array}{l} a_{11}x_1+a_{12}x_2+\ldots+a_{1n}x_n=b_1, \\ a'_{21}x_1+a'_{22}x_2+\ldots+a'_{2n}x_n=b'_2, \\ a_{31}x_1+a_{32}x_2+\ldots+a_{3n}x_n=b_3, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ a_{s1}x_1+a_{s2}x_2+\ldots+a_{sn}x_n=b_s, \end{array}\right\} \tag{4}$$

* También se llama método de reducción. (*Nota del T.*)

donde

$$a'_{2j}=a_{2j}-ca_{1j} \text{ para } j=1, 2, \ldots, n, \ b'_2=b_2-cb_1.$$

Los sistemas de ecuaciones (1) *y* (4) *son equivalentes, es decir, son simultáneamente incompatibles o son simultáneamente compatibles y, en el último caso, poseen las mismas soluciones.* En efecto, sea $k_1, k_2, \ldots, k_n$ una solución arbitraria del sistema (1). Es evidente, que estos números satisfacen a todas las ecuaciones del sistema (4), menos a la segunda. Sin embargo, también satisfacen a la segunda ecuación del sistema (4): es suficiente recordar que esta ecuación se expresa mediante la segunda y la primera de las ecuaciones del sistema (1). Recíprocamente, toda solución del sistema (4) satisface también al sistema (1). En efecto, la segunda ecuación del sistema (1) se obtiene restando de ambos miembros de la segunda ecuación del sistema (4) los miembros correspondientes de la primera ecuación de este sistema, multiplicados por el número $-c$.

Es comprensible que, *si en el sistema* (1) *se efectúan unas cuantas veces las transformaciones del tipo considerado, el sistema obtenido de ecuaciones se mantendrá equivalente al sistema inicial* (1).

Puede ocurrir que después de efectuar tales transformaciones aparezca en nuestro sistema una ecuación, cuyos coeficientes en el primer miembro sean iguales a cero. Si el término independiente de esta ecuación es también igual a cero, la ecuación se satisface con cualesquiera valores de las incógnitas. Por lo tanto, *suprimiendo esta ecuación, llegamos a un sistema de ecuaciones que es equivalente al inicial.* Si el término independiente de la ecuación considerada es diferente de cero, la ecuación no puede ser satisfecha por ninguno de los valores de las incógnitas y, por esto, *el sistema obtenido de ecuaciones, al igual que el sistema inicial equivalente, será incompatible.*

Expongamos ahora el método de Gauss.

Sea dado un sistema arbitrario de ecuaciones lineales (1). Supongamos para precisar que $a_{11}\neq 0$; claro, puede ocurrir que a_{11} sea igual a cero y, entonces, tendríamos que comenzar por cualquier otro coeficiente de la primera ecuación del sistema, diferente de cero.

Transformemos ahora el sistema (1), eliminando la incógnita x_1 de todas las ecuaciones, menos de la primera. Para esto, multipliquemos ambos miembros de la primera ecuación por $\frac{a_{21}}{a_{11}}$ y restémoslos de los miembros correspondientes de la segunda ecuación. Después, multipliquemos ambos miembros de la primera ecuación por $\frac{a_{31}}{a_{11}}$, y restémoslos de los miembros correspondientes de la tercera ecuación, etc., etc.

De este modo, obtendremos un nuevo sistema de s ecuaciones lineales con n incógnitas:

$$\left.\begin{aligned} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 + \ldots + a_{1n}x_n &= b_1, \\ a'_{22}x_2 + a'_{23}x_3 + \ldots + a'_{2n}x_n &= b'_2, \\ a'_{32}x_2 + a'_{33}x_3 + \ldots + a'_{3n}x_n &= b'_3, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots & \\ a'_{s2}x_2 + a'_{s3}x_3 + \ldots + a'_{sn}x_n &= b'_s. \end{aligned}\right\} \quad (5)$$

No tenemos necesidad de escribir explícitamente las expresiones de los coeficientes nuevos a'_{ij} y de los términos independientes nuevos b'_i, mediante los coeficientes y los términos independientes del sistema inicial (1).

Como ya sabemos, el sistema de ecuaciones (5) es equivalente al sistema (1). Transformemos ahora el sistema (5). Pero no tocaremos más la primera ecuación, y *las transformaciones solamente las efectuaremos* con la parte del sistema (5) formada por todas las ecuaciones, menos la primera. Se sobrentiende que entre ellas no hay ecuaciones cuyos coeficientes de los primeros miembros sean iguales a cero: tales ecuaciones las habríamos suprimido, si sus términos independientes fuesen iguales a cero, y en caso contrario, quedaría ya demostrada la incompatibilidad de nuestro sistema. Por lo tanto, hay coeficientes a'_{ij} diferentes de cero; supongamos, para precisar, que $a'_{22} \neq 0$. Transformemos ahora el sistema (5), restando de ambos miembros de la tercera ecuación y de cada una de las siguientes ecuaciones, ambos miembros de la segunda ecuación, multiplicados por

$$\frac{a'_{32}}{a'_{22}}, \quad \frac{a'_{42}}{a'_{22}}, \quad \ldots, \quad \frac{a'_{s2}}{a'_{22}},$$

respectivamente. De este modo, quedará eliminada x_2 de todas las ecuaciones, menos de la primera y de la segunda. Obtendremos el siguiente sistema de ecuaciones, que es equivalente al sistema (5), y por consiguiente, también al sistema (1):

$$\left.\begin{aligned} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 + \ldots + a_{1n}x_n &= b_1, \\ a'_{22}x_2 + a'_{23}x_3 + \ldots + a'_{2n}x_n &= b'_2, \\ a''_{33}x_3 + \ldots + a''_{3n}x_n &= b''_3, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots & \\ a''_{t3}x_3 + \ldots + a''_{tn}x_n &= b''_t. \end{aligned}\right\}$$

Nuestro sistema contiene ahora t ecuaciones, $t \leqslant s$, puesto que, posiblemente, algunas de las ecuaciones hayan sido suprimidas. Es evidente que después de eliminar la incógnita x_1, puede disminuir el

número de ecuaciones del sistema. A continuación, transformaremos solamente aquella parte del sistema obtenido que contiene todas las ecuaciones, menos las dos primeras.

¿Cuándo se terminará este proceso de eliminación consecutiva de las incógnitas?

Si llegamos a un sistema tal, en el que una de sus ecuaciones tenga un término independiente diferente de cero, mientras que todos los coeficientes del primer miembro sean iguales a cero, entonces, como ya sabemos, nuestro sistema inicial será incompatible.

En caso contrario, obtendremos el siguiente sistema de ecuaciones, equivalente al sistema (1):

$$\left.\begin{array}{r} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1,k-1}x_{k-1} + a_{1k}x_k + \ldots + a_{1n}x_n = b_1, \\ a'_{22}x_2 + \ldots + a'_{2,k-1}x_{k-1} + a'_{2k}x_k + \ldots + a'_{2n}x_n = b'_2, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ a^{(k-2)}_{k-1,k-1}x_{k-1} + a^{(k-2)}_{k-1,k}x_k + \ldots + a^{(k-2)}_{k-1,n}x_n = b^{(k-2)}_{k-1}, \\ a^{(k-1)}_{kk}x_k + \ldots + a^{(k-1)}_{kn}x_n = b^{(k-1)}_k. \end{array}\right\} \quad (6)$$

Aquí, $a_{11} \neq 0$, $a'_{22} \neq 0$, $\ldots$, $a^{(k-2)}_{k-1,k-1} \neq 0$, $a^{(k-1)}_{kk} \neq 0$. Señalemos también que $k \leqslant s$ y, evidentemente, $k \leqslant n$.

En este caso, el sistema (1) *es compatible. Es determinado para* $k = n$, *e indeterminado, para* $k < n$.

En efecto, si $k = n$, el sistema (6) tiene la forma

$$\left.\begin{array}{r} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n = b_1, \\ a'_{22}x_2 + \ldots + a'_{2n}x_n = b'_2, \\ \ldots\ldots\ldots\ldots\ldots\ldots \\ a^{(n-1)}_{nn}x_n = b^{(n-1)}_n. \end{array}\right\} \quad (7)$$

De la última ecuación obtenemos un valor absolutamente determinado de la incógnita x_n. Sustituyéndolo en la penúltima ecuación, hallaremos un valor unívocamente determinado de la incógnita x_{n-1}. Continuando de este modo, hallaremos que el sistema (7), y por tanto el sistema (1), poseen solución única, es decir, son compatibles y determinados.

Si $k < n$, tomamos valores numéricos arbitrarios para las incógnitas «independientes» $x_{k+1}, \ldots, x_n$, después de lo cual, avanzando por el sistema (6) de abajo arriba hallaremos, como anteriormente, unos valores unívocamente determinados para las incógnitas x_k, $x_{k-1}, \ldots, x_2, x_1$. Como los valores para las incógnitas independientes se pueden elegir de infinitos modos, el sistema (6) y, por consiguiente, el sistema (1), serán compatibles, pero indeterminados. Es fácil comprobar, que con el método indicado (eligiendo de

todos los modos posibles los valores para las incógnitas independientes), se hallan todas las soluciones del sistema (1).

A primera vista puede parecer que con el método de Gauss el sistema de ecuaciones lineales se puede reducir a otra forma más: la que resulta de agregar al sistema (7) unas cuantas ecuaciones que contengan solamente a la incógnita x_n. Sin embargo, lo que ocurre en realidad es que las transformaciones no se han llevado hasta el fin: como $a_{nn}^{(n-1)} \neq 0$, se puede eliminar la incógnita x_n de todas las ecuaciones, comenzando desde la $(n+1)$-ésima.

Se debe advertir, que la forma «triangular» del sistema de ecuaciones (7), o la forma «trapezoidal» del sistema de ecuaciones (6) (para $k < n$), se obtuvo debido a la suposición de que los coeficientes a_{11}, a'_{22}, etc., etc, eran diferentes de cero. En el caso general, el sistema de ecuaciones a que llegaremos después de realizar hasta el fin el proceso de eliminación de las incógnitas, tomará una forma triangular o trapezoidal sólo después de un cambio debido de la numeración de las incógnitas.

Haciendo un resumen de todo lo expuesto anteriormente, llegamos a la conclusión de que *el método de Gauss se puede aplicar a cualquier sistema de ecuaciones lineales. Además, el sistema será incompatible, si en el proceso de las transformaciones obtenemos una ecuación en la que los coeficientes de las incógnitas son iguales a cero, mientras que el término independiente es diferente de cero; si no nos encontramos con tal ecuación, el sistema será compatible. Un sistema compatible de ecuaciones es determinado, si se reduce a la forma triangular* (7), *e indeterminado, si se reduce a la forma trapezoidal* (6) *siendo* $k < n$.

Apliquemos lo expuesto al caso de un sistema de ecuaciones lineales *homogéneas*, es decir, de ecuaciones, cuyos términos independientes son iguales a cero. Tal sistema siempre es compatible, puesto que posee la *solución nula* (0, 0, . . ., 0). Supongamos que en el sistema considerado, el número de ecuaciones es *menor* que el número de incógnitas. Entonces, este sistema no podrá reducirse a la forma triangular, puesto que en el proceso de transformaciones por el método de Gauss, el número de ecuaciones del sistema sólo puede disminuir pero no aumentar; por consiguiente, éste se reducirá a la forma trapezoidal, es decir, será indeterminado.

En otras palabras: *si en un sistema de ecuaciones lineales homogéneas, el número de ecuaciones es menor que el número de incógnitas, este sistema, además de la solución nula, poseerá también soluciones no nulas*, es decir, soluciones, en las que los valores de ciertas incógnitas (o incluso de todas) serán diferentes de cero; *habrá una infinidad de soluciones de éstas.*

Para la resolución práctica de un sistema de ecuaciones lineales por el método de Gauss, se debe escribir la matriz de los coeficientes

del sistema y agregarle una columna formada por los términos independientes, separada para mayor comodidad por una raya vertical. Todas las transformaciones se deben efectuar con las filas de esta matriz «ampliada».

Ejemplos. 1. Resolver el sistema

$$\left.\begin{aligned} x_1+2x_2+5x_3&=-9,\\ x_1-x_2+3x_3&=2,\\ 3x_1-6x_2-x_3&=25. \end{aligned}\right\}$$

Efectuemos las transformaciones en la matriz ampliada del sistema:

$$\left(\begin{array}{rrr|r} 1 & 2 & 5 & -9\\ 1 & -1 & 3 & 2\\ 3 & -6 & -1 & 25 \end{array}\right) \to \left(\begin{array}{rrr|r} 1 & 2 & 5 & -9\\ 0 & -3 & -2 & 11\\ 0 & -12 & -16 & 52 \end{array}\right) \to \left(\begin{array}{rrr|r} 1 & 2 & 5 & -9\\ 0 & -3 & -2 & 11\\ 0 & 0 & -8 & 8 \end{array}\right)$$

Por consiguiente, llegamos al siguiente sistema de ecuaciones:

$$\left.\begin{aligned} x_1+2x_2+5x_3&=-9,\\ -3x_2-2x_3&=11,\\ -8x_3&=8, \end{aligned}\right\}$$

que posee la solución única:

$$x_1=2,\quad x_2=-3,\quad x_3=-1.$$

Por consiguiente, el sistema inicial es determinado.

2. Resolver el sistema

$$\left.\begin{aligned} x_1-5x_2-8x_3+x_4&=3,\\ 3x_1+x_2-3x_3-5x_4&=1,\\ x_1-7x_3+2x_4&=-5,\\ 11x_2+20x_3-9x_4&=2. \end{aligned}\right\}$$

Transformemos la matriz ampliada del sistema:

$$\left(\begin{array}{rrrr|r} 1 & -5 & -8 & 1 & 3\\ 3 & 1 & -3 & -5 & 1\\ 1 & 0 & -7 & 2 & -5\\ 0 & 11 & 20 & -9 & 2 \end{array}\right) \to \left(\begin{array}{rrrr|r} 1 & -5 & 8 & 1 & 3\\ 0 & 16 & 21 & -8 & -8\\ 0 & 5 & 1 & 1 & -8\\ 0 & 11 & 20 & -9 & 2 \end{array}\right) \to$$

$$\to \left(\begin{array}{rrrr|r} 1 & -5 & -8 & 1 & 3\\ 0 & -89 & 0 & -29 & 160\\ 0 & 5 & 1 & 1 & -8\\ 0 & -89 & 0 & -29 & 162 \end{array}\right) \to \left(\begin{array}{rrrr|r} 1 & -5 & -8 & 1 & 3\\ 0 & -89 & 0 & -29 & 160\\ 0 & 5 & 1 & 1 & -8\\ 0 & 0 & 0 & 0 & 2 \end{array}\right)$$

Hemos llegado a un sistema que contiene la ecuación $0=2$. Por consiguiente, el sistema inicial es incompatible.

3. Resolver el sistema

$$\begin{aligned} 4x_1+x_2-3x_3-x_4&=0,\\ 2x_1+3x_2+x_3-5x_4&=0,\\ x_1-2x_2-2x_3+3x_4&=0. \end{aligned}$$

Este es un sistema de ecuaciones homogéneas, donde el número de ecuaciones es menor que el número de incógnitas; por lo tanto, tiene que ser indeterminado. Como todos los términos independientes son iguales a cero, vamos a trans-

formar solamente la matriz de los coeficientes del sistema:

$$\begin{pmatrix} 4 & 1 & -3 & -1 \\ 2 & 3 & 1 & -5 \\ 1 & -2 & -2 & 3 \end{pmatrix} \to \begin{pmatrix} 0 & 9 & 5 & -13 \\ 0 & 7 & 5 & -11 \\ 1 & -2 & -2 & 3 \end{pmatrix} \to \begin{pmatrix} 0 & 2 & 0 & -2 \\ 0 & 7 & 5 & -11 \\ 1 & -2 & -2 & 3 \end{pmatrix}.$$

Hemos obtenido el sistema de ecuaciones:

$$\left.\begin{aligned} 2x_2 \qquad - 2x_4 &= 0, \\ 7x_2 + 5x_3 - 11x_4 &= 0, \\ x_1 - 2x_2 - 2x_3 + 3x_4 &= 0. \end{aligned}\right\}$$

Cualquiera de las incógnitas x_2 y x_4, se puede tomar como incógnita independiente. Sea $x_4 = \alpha$; entonces, de la primera ecuación se deduce que $x_2 = \alpha$; de la segunda ecuación obtenemos, $x_3 = \frac{4}{5}\alpha$; y, por fin, de la tercera ecuación, $x_1 = \frac{3}{5}\alpha$. Por lo tanto, la forma general de las soluciones del sistema de ecuaciones dado es:

$$\frac{3}{5}\alpha, \ \alpha, \ \frac{4}{5}\alpha, \ \alpha.$$

§ 2. Determinantes de segundo y tercer orden

El método de resolución de los sistemas de ecuaciones lineales, expuesto en el párrafo anterior, es muy sencillo y requiere la realización de cálculos de un mismo tipo, que fácilmente se efectúan en las máquinas calculadoras. Sin embargo, su defecto esencial consiste en que no da la posibilidad de formular las condiciones de compatibilidad o de determinabilidad de un sistema mediante sus coeficientes y términos independientes. Por otra parte, incluso en el caso de un sistema determinado, con este método no se pueden hallar fórmulas para expresar la solución del sistema mediante sus coeficientes y términos independientes. Sin embargo, estos sistemas encuentran aplicación en diversas cuestiones teóricas y, en particular, en las investigaciones geométricas. De aquí la necesidad de desarrollar la teoría de los sistemas de ecuaciones lineales con otros métodos más profundos. El caso general va a ser estudiado en el capítulo siguiente, mientras que el contenido del presente capítulo está dedicado al estudio de los sistemas determinados que tienen igual número de ecuaciones y de incógnitas. Comenzaremos por los sistemas con dos y tres incógnitas, ya estudiados en el álgebra elemental.

Sea dado un *sistema de dos ecuaciones lineales con dos incógnitas*

$$\left.\begin{aligned} a_{11}x_1 + a_{12}x_2 &= b_1, \\ a_{21}x_1 + a_{22}x_2 &= b_2, \end{aligned}\right\} \tag{1}$$

cuyos coeficientes forman una matriz cuadrada de segundo orden

$$\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}. \qquad (2)$$

Aplicando al sistema (1), el método de igualación de los coeficientes obtenemos:

$$(a_{11}a_{22}-a_{12}a_{21})\,x_1 = b_1a_{22}-a_{12}b_2,$$
$$(a_{11}a_{22}-a_{12}a_{21})\,x_2 = a_{11}b_2-b_1a_{21}.$$

Supongamos que $a_{11}a_{22}-a_{12}a_{21}\neq 0$. Entonces,

$$x_1 = \frac{b_1a_{22}-a_{12}b_2}{a_{11}a_{22}-a_{12}a_{21}}, \quad x_2 = \frac{a_{11}b_2-b_1a_{21}}{a_{11}a_{22}-a_{12}a_{21}}. \qquad (3)$$

Sustituyendo en las ecuaciones (1) los valores obtenidos de las incógnitas, es fácil comprobar que (3) es solución del sistema (1); el problema de la unicidad de esta solución se estudiará en el § 7.

El común denominador de los valores de las incógnitas (3) está expresado sencillamente por los elementos de la matriz (2), o sea, es precisamente igual al producto de los elementos de la diagonal principal menos el producto de los elementos de la segunda diagonal. Este número se llama *determinante* de la matriz (2). Se suele decir que es un *determinante de segundo orden*, puesto que la matriz (2) es de segundo orden. Para designar el determinante de la matriz (2), se emplea la siguiente notación: se escribe la matriz (2), pero en lugar de los paréntesis se ponen unas barras verticales; de este modo

$$\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11}a_{22}-a_{12}a_{21}. \qquad (4)$$

Ejemplos.

1) $$\begin{vmatrix} 3 & 7 \\ 1 & 4 \end{vmatrix} = 3\cdot 4-7\cdot 1 = 5;$$

2) $$\begin{vmatrix} 1 & -2 \\ 3 & 5 \end{vmatrix} = 1\cdot 5-(-2)\cdot 3 = 11$$

Es menester subrayar otra vez más que, mientras la matriz representa una tabla de números, el determinante es un número completamente determinado por la matriz cuadrada. Señalemos que los productos $a_{11}a_{22}$ y $a_{12}a_{21}$, se llaman *términos* del determinante de segundo orden.

Los numeradores de las expresiones (3) tienen la misma forma que el denominador, o sea, también son determinantes de segundo orden: el numerador de la expresión para x_1 es el determinante de una matriz, que se obtiene de la matriz (2) sustituyendo su primera columna por la columna de los términos independientes del sistema

(1); el numerador de la expresión para x_2 es el determinante de una matriz, que se obtiene de la matriz (2) por la misma sustitución de su segunda columna. Las fórmulas (3) se pueden escribir ahora en la forma siguiente:

$$x_1 = \frac{\begin{vmatrix} b_1 & a_{12} \\ b_2 & a_{22} \end{vmatrix}}{\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}}, \quad x_2 = \frac{\begin{vmatrix} a_{11} & b_1 \\ a_{21} & b_2 \end{vmatrix}}{\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}}. \tag{5}$$

Esta regla de resolución de un sistema de dos ecuaciones lineales con dos incógnitas (denominada *regla de Cramer*) se expresa del modo siguiente:

Si el determinante (4) *de los coeficientes del sistema de ecuaciones* (1) *es diferente de cero, la solución del sistema* (1) *se obtiene tomando por valores de las incógnitas las fracciones, cuyo común denominador es el determinante* (4) *y cuyo numerador, para la incógnita* x_i ($i = 1, 2$), *es el determinante que se obtiene sustituyendo en el determinante* (4) *la columna i-ésima* (*o sea, la columna de los coeficientes de la incógnita buscada*) *por la columna de los términos independientes del sistema* (1)*.

Ejemplo. Resolver el sistema

$$\left.\begin{aligned} 2x_1 + x_2 &= 7, \\ x_1 - 3x_2 &= -2. \end{aligned}\right\}$$

El determinante de los coeficientes es

$$d = \begin{vmatrix} 2 & 1 \\ 1 & -3 \end{vmatrix} = -7;$$

es decir, éste es diferente de cero, por lo cual, se puede aplicar la regla de Cramer al sistema.

Los numeradores para las incógnitas son los determinantes:

$$d_1 = \begin{vmatrix} 7 & 1 \\ -2 & -3 \end{vmatrix} = -19, \quad d_2 = \begin{vmatrix} 2 & 7 \\ 1 & -2 \end{vmatrix} = -11.$$

Por lo tanto, la solución de nuestro sistema es:

$$x_1 = \frac{d_1}{d} = \frac{19}{7}, \quad x_2 = \frac{d_2}{d} = \frac{11}{7}.$$

La introducción de los determinantes de segundo orden no aporta simplificaciones esenciales en la resolución de un sistema de dos ecuaciones lineales con dos incógnitas. Sin embargo, los métodos análogos para el caso de *sistemas de tres ecuaciones lineales con tres*

* En este enunciado, para abreviar, se habla de la sustitución de las columnas «en el determinante». A continuación, se va a hablar de un modo semejante, si resulta conveniente, de las filas y columnas del determinante, de sus elementos, diagonales, etc.

incógnitas resultan ya prácticamente útiles. Sea dado un sistema

$$\left.\begin{aligned} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 = b_1, \\ a_{21}x_1 + a_{22}x_2 + a_{23}x_3 = b_2, \\ a_{31}x_1 + a_{32}x_2 + a_{33}x_3 = b_3 \end{aligned}\right\} \qquad (6)$$

con la matriz de los coeficientes

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}. \qquad (7)$$

Fácilmente se comprueba que, si multiplicamos ambos miembros de la primera de las ecuaciones (6) por $a_{22}a_{33} - a_{23}a_{32}$, ambos miembros de la segunda ecuación por $a_{13}a_{32} - a_{12}a_{33}$, ambos miembros

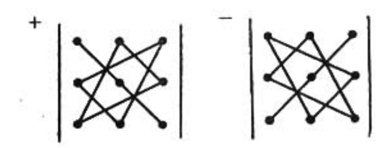


Fig 1.

de la tercera ecuación por $a_{12}a_{23} - a_{13}a_{22}$, y después sumamos estas tres ecuaciones, los coeficientes de x_2 y x_3 resultarán iguales a cero, es decir, que estas incógnitas se eliminarán simultáneamente. De este modo, obtenemos la igualdad

$$(a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} - a_{13}a_{22}a_{31} - a_{12}a_{21}a_{33} - a_{11}a_{23}a_{32})\, x_1 =$$
$$= b_1a_{22}a_{33} + a_{12}a_{23}b_3 + a_{13}b_2a_{32} - a_{13}a_{22}b_3 - a_{12}b_2a_{33} - b_1a_{23}a_{32}. \qquad (8)$$

El coeficiente de x_1 en esta igualdad se llama *determinante de tercer orden*, correspondiente a la matriz (7). Para escribirlo, se emplean los mismos símbolos que en el caso de los determinantes de segundo orden; por lo tanto,

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} - a_{13}a_{22}a_{31} - a_{12}a_{21}a_{33} - a_{11}a_{23}a_{32} \qquad (9)$$

A pesar de que la expresión del determinante de tercer orden es bastante complicada, la ley de su formación con los elementos de la matriz (7) es muy sencilla. En efecto, uno de los tres términos del determinante que figuran en la expresión (9) con el signo más es el producto de los elementos de la diagonal principal; cada uno de los

otros dos, es el producto de los elementos situados en la paralela a esta diagonal, por el elemento situado en el ángulo opuesto de la matriz. Los términos que figuran en (9) con signo menos, se forman del mismo modo, pero con respecto a la segunda diagonal. De este modo, obtenemos un método de cálculo de los determinantes de tercer orden, que conduce (teniendo cierta práctica) a un rápido resultado. En la fig. 1, en el esquema de la izquierda, se señala la regla para el cálculo de los términos positivos del determinante de tercer orden, y en el de la derecha, la regla para el cálculo de sus términos negativos.

Ejemplos.

$$1)\ \begin{vmatrix} 2 & 1 & 2 \\ -4 & 3 & 1 \\ 2 & 3 & 5 \end{vmatrix} = 2\cdot 3\cdot 5 + 1\cdot 1\cdot 2 + 2\cdot(-4)\cdot 3 - 2\cdot 3\cdot 2 - 1\cdot(-4)\cdot 5 - 2\cdot 1\cdot 3 = 30 + 2 - 24 - 12 + 20 - 6 = 10.$$

$$2)\ \begin{vmatrix} 1 & 0 & -5 \\ -2 & 3 & 2 \\ 2 & -2 & 0 \end{vmatrix} = 1\cdot 3\cdot 0 + 0\cdot 2\cdot 1 + (-5)\cdot(-2)\cdot(-2) - (-5)\cdot 3\cdot 1 - 0\cdot(-2)\cdot 0 - 1\cdot 2\cdot(-2) = -20 + 15 + 4 = -1.$$

El segundo miembro de la igualdad (8) es también un determinante de tercer orden: es precisamente el determinante de la matriz que se obtiene de la matriz (7), sustituyendo su primera columna por la columna de los términos independientes del sistema (6). Si designamos con la letra d el determinante (9) y con el símbolo d_j ($j = 1, 2, 3$), el determinante que se obtiene de este último, al sustituir su j-ésima columna por la columna de los términos independientes del sistema (6), la igualdad (8) toma la forma $dx_1 = d_1$, de donde, para $d \neq 0$, se deduce que

$$x_1 = \frac{d_1}{d}\ . \qquad (10)$$

Del mismo modo, multiplicando las ecuaciones (6) por los números $a_{23}a_{31} - a_{21}a_{33}$, $a_{11}a_{33} - a_{13}a_{31}$, $a_{13}a_{21} - a_{11}a_{23}$, respectivamente, obtenemos para x_2 la siguiente expresión (siendo $\neq 0$):

$$x_2 = \frac{d_2}{d} \qquad (11)$$

Finalmente, multiplicando estas ceuaciones por $a_{21}a_{32} - a_{22}a_{31}$, $a_{12}a_{31} - a_{11}a_{32}$, $a_{11}a_{22} - a_{12}a_{21}$, respectivamente, llegamos a la siguiente expresión para x_3:

$$x_3 = \frac{d_3}{d}\ . \qquad (12)$$

Sustituyendo las expresiones (10), (11) y (12) en la ecuación (6) (se sobrentiende que los determinantes d y todos los d_j están escritos en forma desarrollada), después de unos cálculos bastante complicados, pero asequibles, vemos que se satisfacen todas estas ecuaciones, es decir, que los números (10) (11) y (12) son la solución del sistema. Por lo tanto, *si el determinante de los coeficientes de un sistema de tres ecuaciones lineales con tres incógnitas es diferente de cero, su solución se puede hallar por la regla de Cramer, formulada igualmente que en el caso de un sistema de dos ecuaciones.* En el § 7, el lector hallará, para un caso más general, otra demostración de esta afirmación (que no se basa en los cálculos omitidos), y también la demostración de la unicidad de la solución (10) (11) y (12) del sistema (6).

Ejemplo. Resolver el sistema

$$\left.\begin{aligned} 2x_1 - x_2 + x_3 &= 0, \\ 3x_1 + 2x_2 - 5x_3 &= 1, \\ x_1 + 3x_2 - 2x_3 &= 4. \end{aligned}\right\}$$

El determinante de los coeficientes del sistema es diferente de cero:

$$d = \begin{vmatrix} 2 & -1 & 1 \\ 3 & 2 & -5 \\ 1 & 3 & -2 \end{vmatrix} = 28,$$

por eso, se le puede aplicar la regla de Cramer. Los numeradores para las incógnitas son los determinantes

$$d_1 = \begin{vmatrix} 0 & -1 & 1 \\ 1 & 2 & -5 \\ 4 & 3 & 2 \end{vmatrix} = 13, \quad d_2 = \begin{vmatrix} 2 & 0 & 1 \\ 3 & 1 & -5 \\ 1 & 4 & -2 \end{vmatrix} = 47,$$

$$d_3 = \begin{vmatrix} 2 & -1 & 0 \\ 3 & 2 & 1 \\ 1 & 3 & 4 \end{vmatrix} = 21,$$

o sea, que la solución del sistema es el sistema de números

$$x_1 = \frac{13}{28}, \quad x_2 = \frac{47}{28}, \quad x_3 = \frac{21}{28} = \frac{3}{4}.$$

§ 3. Permutaciones y sustituciones

Para definir y estudiar los determinantes de orden n, necesitamos unos cuantos conceptos y datos referentes a los conjuntos finitos. Sea dado un conjunto finito M, compuesto de n elementos. Estos se pueden numerar, empleando para ello los primeros n números naturales 1, 2, . . ., n. Como en las cuestiones que nos interesan, las propiedades individuales de los elementos del conjunto M no van

a jugar ningún papel, supondremos simplemente que los mismos números 1, 2, . . ., n, representan a los elementos del conjunto M.

Además de la ordenación normal de los números 1, 2, . . ., n, éstos pueden ser ordenados de muchos modos. Así, los números 1, 2, 3, 4 se pueden ordenar también de los modos siguientes: 3,1, 2, 4, o bien, 2, 4, 1, 3, etc. Toda disposición de los números 1, 2, . . . ,n en un orden determinado, se llama *permutación* de n números (o de n símbolos).

El número de permutaciones diversas de n símbolos es igual al producto $1 \cdot 2 \ldots n$, designado por $n!$ (se lee: «factorial de n»). En efecto, la forma general de una permutación de n símbolos es $i_1, i_2, \ldots, i_n$, donde cada uno de los símbolos i_s representa uno de los números 1, 2, . . ., n; además, ninguno de estos números se repite. En calidad de i_1 se puede tomar cualquiera de los números 1, 2, . . ., n; esto ofrece n diferentes posibilidades. Sin embargo, si se ha elegido ya i_1, en calidad de i_2 se puede tomar solamente uno de los $n-1$ números restantes, es decir, que el número de modos de elección de los símbolos i_1 y i_2 es igual al producto $n(n-1)$, etc.

Por lo tanto, el número de permutaciones de n símbolos, para $n = 2$, es igual a $2! = 2$ (las permutaciones 12 y 21; en los ejemplos donde $n \leqslant 9$, no separaremos con comas los símbolos que se permutan); para $n = 3$, este número es igual a $3! = 6$, para $n = 4$, es igual a $4! = 24$. A continuación, con el aumento de n, el número de permutaciones crece extraordinariamente; así, para $n = 5$, es igual a $5! = 120$, y para $n = 10$, es ya igual a 3 628 800.

Si en una permutación cambiamos de lugar dos símbolos cualesquiera (no necesariamente situados uno al lado del otro), permaneciendo todos los demás en sus sitios, obtenemos, evidentemente, una nueva permutación. Esta transformación de la permutación se denomina *trasposición*.

Todas las $n!$ permutaciones de n símbolos se pueden colocar en un orden tal, que cada permutación siguiente se obtenga de la anterior mediante una trasposición, pudiendo además comenzar por cualquiera de ellas.

Esta afirmación es justa para $n = 2$: si se pide empezar por la permutación 12, la disposición buscada es 12, 21; si se pide empezar por la permutación 21, la disposición es 21, 12. Supongamos, que nuestra afirmación ya está demostrada para $n - 1$, y que queremos demostrarla para n. Supongamos, además, que tenemos que comenzar con la permutación

$$i_1, i_2, \ldots, i_n. \tag{1}$$

Consideremos todas las permutaciones de n símbolos en las que i_1 ocupa el primer lugar. En total, resultan $(n-1)!$ permutaciones. Según la tesis del teorema, éstas pueden ser ordenadas del modo indi-

cado y, además, comenzando por la permutación (1), puesto que, en realidad, esto se reduce a la ordenación de todas las permutaciones de $n - 1$ símbolos y, por la suposición inductiva, se puede comenzar por cualquier permutación y, en particular, por la permutación $i_2, \ldots, i_n$. En la última de las permutaciones de n símbolos obtenidas de este modo, efectuamos una trasposición del símbolo i_1 con cualquier otro símbolo, por ejemplo, con i_2, y comenzando con la permutación nueva obtenida, ordenamos del modo necesario todas las permutaciones en las que i_2 ocupa el primer lugar, etc. Es evidente, que de este modo se pueden obtener todas las permutaciones de n símbolos.

De este teorema se deduce que *de cualquier permutación de n símbolos se puede pasar a cualquier otra permutación de los mismos símbolos, mediante unas cuantas trasposiciones.*

Se dice que, en una permutación dada, los números i y j forman una *inversión*, si $i > j$, pero en esta permutación, i está antes que j. Una permutación se llama *par*, si sus símbolos forman un número par de inversiones, e *impar*, en el caso contrario. Así, la permutación $1, 2, \ldots, n$ es par para cualquier n, puesto que en ella el número de inversiones es igual a cero. La permutación 451362 ($n = 6$) contiene 8 inversiones y, por consiguiente, es par; la permutación 38524671 ($n = 8$) contiene 15 inversiones y, por consiguiente, es impar.

Toda trasposición cambia la paridad de la permutación.

Para demostrar este importante teorema, consideremos primero el caso en que los símbolos i y j que se trasponen estén uno al lado del otro, es decir, que la permutación tiene la forma..., $i, j, \ldots$, donde los puntos sustituyen a los símbolos que no se alteran con la trasposición. La trasposición convierte a nuestra permutación en la permutación..., $j, i, \ldots$; se comprende, además, que en ambas permutaciones, cada uno de los símbolos i, j, forma unas mismas inversiones con los símbolos que se mantienen en el sitio. Si antes los símbolos i y j no formaban inversión, en la nueva permutación aparece una nueva inversión, o sea, el número de inversiones aumenta en una unidad; si antes los símbolos i y j formaban inversión, ésta ahora desaparece, o sea, el número de inversiones disminuye en una unidad. En ambos casos, cambia la paridad de la permutación.

Supongamos ahora, que entre los símbolos i y j que se trasponen hay intercalados s símbolos, $s > 0$, es decir, que la permutación tiene la forma

$$\ldots, i, k_1, k_2, \ldots, k_s, j, \ldots \tag{2}$$

Se puede obtener la trasposición de los símbolos i y j como resultado de la ejecución consecutiva de $2s + 1$ trasposiciones de elementos vecinos. Estas son, precisamente, las trasposiciones que permutan

los símbolos i y k_1, a continuación, i (que ya ocupa el lugar del símbolo k_1) y k_2, etc, hasta que i llegue a ocupar el lugar del símbolo k_s. Después de estas s trasposiciones viene la trasposición que permuta los símbolos i y j, y luego, s trasposiciones del símbolo j con todos los k, a consecuencia de lo cual j ocupa el lugar del símbolo i, mientras que los símbolos k vuelven a sus lugares antiguos. Por lo tanto, la paridad de la permutación fue cambiada un número impar de veces, y por esto, la permutación (2) y

$$\ldots,\ j,\ k_1,\ k_2,\ \ldots,\ k_s i,\ \ldots \tag{3}$$

tienen diferente paridad.

Para $n \geqslant 2$, *el número de permutaciones pares de n símbolos es igual al número de permutaciones impares, es decir, es igual a* $\frac{1}{2}\, n!$

En efecto, basándonos en lo demostrado anteriormente, ordenemos todas las permutaciones de n símbolos de tal modo, que cada una de ellas se obtenga de la anterior mediante una trasposición. Las permutaciones vecinas tendrán entonces paridad contraria, es decir, las permutaciones estarán colocadas de tal manera, que las permutaciones pares e impares se alternarán. Nuestra afirmación se deduce ahora de la observación evidente que para $n \geqslant 2$, el número $n!$ es par.

Introduzcamos ahora un nuevo concepto, el *de sustitución de grado n*. Escribamos, una debajo de otra, dos permutaciones de n símbolos, colocando entre paréntesis las dos filas obtenidas; por ejemplo, para $n = 5$:

$$\begin{pmatrix} 3 & 5 & 1 & 4 & 2 \\ 5 & 2 & 3 & 4 & 1 \end{pmatrix}. \tag{4}$$

En este ejemplo*, bajo el número 3 figura el número 5, bajo el número 5, el número 2, etc. Diremos que el número 3 *se sustituye* por el 5 (o también, que la sustitución transporta el 3 sobre el 5); el número 5, por el 2; el número 1, por el 3; número 4, por el 4 (o que *se queda en el sitio*); y, por fin, el número 2, por el 1. Por lo tanto, dos permutaciones, escritas una bajo la otra en la forma (4), determinan una *aplicación biyectiva*** *del conjunto* de los primeros cinco números

---

* Por su aspecto, se parece a una matriz de dos filas y 5 columnas, pero tiene un significado totalmente distinto.

** A continuación se empleará frecuentemente la siguiente terminología, admitida en la teoría de conjuntos.

Sean dados dos conjuntos M y N (finitos o infinitos) de elementos de cualquier naturaleza.

Si de un modo determinado, a cada elemento x de M (con la notación $x \in M$ se denota que el elemento x pertenece a M) se pone en correspondencia un elemento y de N ($y \in N$), y sólo uno, se dice que se ha definido una apli-

naturales sobre sí mismo, es decir, una aplicación que, a cada uno de los números naturales 1, 2, 3, 4, 5, pone en correspondencia uno de estos mismos números naturales, y que, a diversos números pone en correspondencia números diferentes. Como en total hay cinco números de éstos, o sea, es un conjunto finito, a *cada uno* de estos números corresponderá uno de los números 1, 2, 3, 4, 5, es decir, precisamente el número por el que «se sustituye».

Está claro, que la aplicación biyectiva del conjunto de los cinco primeros números naturales que hemos obtenido mediante (4), se podría haber obtenido también escribiendo, una bajo la otra, otros pares de permutaciones de los cinco símbolos. Estas expresiones se obtienen de (4) mediante unas cuantas trasposiciones de las columnas; tales son, por ejemplo,

$$\begin{pmatrix} 2 & 1 & 5 & 3 & 4 \\ 1 & 3 & 2 & 5 & 4 \end{pmatrix}, \quad \begin{pmatrix} 1 & 5 & 2 & 4 & 3 \\ 3 & 2 & 1 & 4 & 5 \end{pmatrix}, \quad \begin{pmatrix} 2 & 5 & 1 & 4 & 3 \\ 1 & 2 & 3 & 4 & 5 \end{pmatrix}. \qquad (5)$$

En todas estas expresiones, 3 se sustituye por 5, 5 por 2, etc.

De modo análogo, dos permutaciones de n símbolos, escritas una bajo la otra, determinan una aplicación biyectiva del conjunto de los primeros n números naturales sobre sí mismo. Toda aplicación biyectiva A del conjunto de los primeros n números naturales sobre sí mismo, se llama *sustitución de grado n*. Es evidente, que toda sustitución A se puede expresar mediante dos permutaciones, escritas una bajo la otra

$$A = \begin{pmatrix} i_1, & i_2, & \dots & i_n \\ \alpha_{i_1}, & \alpha_{i_2}, & \dots & \alpha_{i_n} \end{pmatrix}; \qquad (6)$$

aquí, mediante α_i se denota el número que en la sustitución A sustituye al número i, $i = 1, 2, \dots, n$.

La sustitución A posee una multitud de expresiones de la forma (6). Así, pues, (4) y (5) son diversas expresiones de una misma sustitución de 5° grado.

Se puede pasar de una expresión de la sustitución A a otra, realizando unas cuantas trasposiciones de las columnas. También se puede obtener una expresión de la forma (6), en la que figure, en la fila superior (o inferior), una permutación prefijada de n símbolos. En

---

cación (o una representación) de M en N. El elemento y se llama en este caso imagen o representación de x.

Si todo elemento de N es imagen de al menos un elemento de M, se dice que se tiene una aplicación de M sobre N (pudiendo ser pluriunívoca, cuando varios elementos de M tienen una misma imagen). En este caso, la aplicación se llama exhaustiva (o sobreyectiva). Si distintos elementos de M tienen distintas imágenes, la aplicación es inyectiva. Una aplicación exhaustiva e inyectiva se llama biyectiva (también suele decirse que entre M y N se ha establecido una correspondencia biunívoca). En este caso, cada elemento $y \in N$ es imagen de un sólo elemento $x \in M$. (*Nota del T.*).

particular, toda sustitución A de grado n se puede expresar en la forma

$$A = \begin{pmatrix} 1 & 2 & \dots & n \\ \alpha_1 & \alpha_2 & \dots & \alpha_n \end{pmatrix}, \tag{7}$$

o sea, con la ordenación natural de los números en la fila superior. Escribiéndolas de este modo, las diversas sustituciones se diferenciarán unas de otras por las permutaciones que figuran en las filas inferiores, con lo que llegamos a la conclusión de que el ***número de sustituciones de grado n es igual al número de permutaciones de n símbolos, es decir, es igual a n!.***

Un ejemplo de sustitución de grado n es la ***sustitución unidad*** (*o idéntica*)

$$E = \begin{pmatrix} 1 & 2 & \dots & n \\ 1 & 2 & \dots & n \end{pmatrix},$$

en la que todos los símbolos permanecen en su sitio.

Señalemos que, en la expresión (6) de la sustitución A, las filas superior e inferior desempeñan papeles diferentes y que, por lo general, cambiándolas de sitio, obtenemos una sustitución diferente. Así pues, las sustituciones de 4° grado

$$\begin{pmatrix} 2 & 1 & 4 & 3 \\ 4 & 3 & 1 & 2 \end{pmatrix} \text{ y } \begin{pmatrix} 4 & 3 & 1 & 2 \\ 2 & 1 & 4 & 3 \end{pmatrix}$$

son distintas: en la primera, el número 2 se sustituye por el 4, mientras que en la segunda, por el 3.

Tomemos una expresión arbitraria (6) de una sustitución A de grado n. Las permutaciones que forman las filas superior e inferior de esta expresión pueden ser de igual paridad o de paridad contraria. Como ya sabemos, el paso a cualquier otra expresión de la sustitución A se puede realizar mediante la ejecución consecutiva de unas cuantas trasposiciones en la fila superior y las trasposiciones correspondientes en la fila inferior. Por otra parte, al efectuar una trasposición en la fila superior de la expresión (6) y una trasposición de los elementos correspondientes en la fila inferior, las paridades de ambas permutaciones cambian simultáneamente, manteniéndose así la coincidencia o la contrariedad de estas paridades. De aquí se deduce que, *en todas las expresiones de la sustitución, las paridades de las filas superior e inferior coinciden, o bien, en todas estas expresiones, las paridades son contrarias.* En el primer caso, se dice que la sustitución A es *par*, en el segundo, que es *impar*. En particular, la sustitución unidad es par.

Si la sustitución A está escrita en la forma (7), es decir, que en la fila superior figura la permutación par $1, 2, \dots, n$, entonces, la paridad

de la sustitución A se determinará por la paridad de la permutación $\alpha_1, \alpha_2, \ldots, \alpha_n$ que figura en la fila inferior. De esto se deduce que el *número de las permutaciones pares de grado n es igual al número de las impares, es decir, es igual a* $\frac{1}{2}n!$.

A la definición de la paridad de las sustituciones se le puede dar otra forma un poco diferente. Si en la expresión (6), las paridades de ambas filas coinciden, el número de inversiones, o es par en ambas filas, o es impar en las dos, es decir, el número total de inversiones en las dos filas de la expresión (6) es par; si las paridades de las filas de la expresión (6) son contrarias, el número total de inversiones en estas dos filas es impar. Por lo tanto, *la sustitución A será par, si el número total de inversiones en las dos filas de cualquiera de sus expresiones es par, e impar, en el caso contrario.*

Ejemplo. Sea dada la sustitución de quinto grado

$$\begin{pmatrix} 3 & 1 & 4 & 5 & 2 \\ 2 & 5 & 4 & 3 & 1 \end{pmatrix}$$

En la fila superior hay 4 inversiones y en la inferior 7. El número total de inversiones en las dos filas es igual a 11 y, por consiguiente, la sustitución es impar.

Escribamos esta sustitución en la forma

$$\begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 5 & 1 & 2 & 4 & 3 \end{pmatrix} .$$

El número de inversiones en la fila superior es 0 y en la inferior es 5, es decir, el número total de inversiones es de nuevo impar. Vemos, pues, que en diversas expresiones de la sustitución se conserva la paridad, pero no el mismo número total de inversiones.

Queremos señalar ahora otras formas de definición de la paridad de las sustituciones que son equivalentes a las expuestas anteriormente*. Para este fin, definiremos el *producto de sustituciones* (que es también de particular interés). Como ya sabemos, la sustitución de grado n es una aplicación biyectiva del conjunto de los números $1, 2, \ldots, n$, sobre sí mismo. El resultado de la realización consecutiva de dos aplicaciones biyectivas del conjunto $1, 2, \ldots, n$ sobre sí mismo es, evidentemente, una nueva aplicación biyectiva de este conjunto sobre sí mismo, es decir, la realización consecutiva de dos sustituciones de grado n da lugar a otra sustitución tercera de grado n, absolutamente determinada. Esta última se llama *producto* de la primera de las sustituciones dadas por la segunda. Así, sean dadas las sustituciones de cuarto grado

$$A = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 3 & 1 & 4 & 2 \end{pmatrix}, \quad B = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 1 & 3 & 4 & 2 \end{pmatrix}.$$

* Estas se necesitarán solamente en el capítulo 14 y por esto, en la primera lectura, este material se puede omitir.

entonces

$$AB=\begin{pmatrix}1&2&3&4\\4&1&2&3\end{pmatrix}.$$

En efecto, en la sustitución A el símbolo 1 se sustituye por el 3, pero en la sustitución B el símbolo 3 se sustituye por el 4, por lo tanto, en la sustitución AB el símbolo 1 se sustituye por el 4, etc.

Solamente se pueden multiplicar las sustituciones de un mismo grado. *Para $n \geqslant 3$, el producto de las sustituciones de grado n no es conmutativo.* En efecto, para las sustituciones A y B consideradas anteriormente, el producto BA tiene la forma

$$BA=\begin{pmatrix}1&2&3&4\\3&4&2&1\end{pmatrix}$$

o sea, la sustitución BA es diferente de la sustitución AB. Para todas las n ($n \geqslant 3$) se pueden mostrar ejemplos de este tipo, a pesar de que para algunos pares de sustituciones se pueda cumplir eventualmente la ley conmutativa.

El producto de las sustituciones es asociativo, es decir, que se puede hablar del producto de un número finito cualquiera de sustituciones de grado n, tomados (en vista de que no se cumple la ley conmutativa) en un orden determinado. En efecto, sean dadas las sustituciones A, B y C. Supongamos que en la sustitución A el símbolo i_1, $1 \leqslant i_1 \leqslant n$ se sustituye por el símbolo i_2; en la sustitución B, el símbolo i_2 se sustituye por el símbolo i_3 y en la sustitución C, este último se sustituye por el símbolo i_4. Entonces, en la sustitución AB, el símbolo i_1 se sustituirá por el i_3, en la sustitución BC, el símbolo i_2 se sustituirá por el i_4. Por consiguiente, en la sustitución $(AB)C$, así como en la sustitución $A\,(BC)$, el símbolo i_1 se sustituirá por el símbolo i_4.

Es evidente, que el *producto de cualquier sustitución A por la sustitución unidad E, y también el producto de E por A, son iguales a A*:

$$AE = EA = A.$$

Finalmente, denominaremos *inversa* de la sustitución A a una sustitución del mismo grado A^{-1}, que cumpla las condiciones

$$AA^{-1} = A^{-1}A = E.$$

Fácilmente se observa que la inversa de la sustitución

$$A=\begin{pmatrix}1&2&\dots&n\\\alpha_1&\alpha_2&&\alpha_n\end{pmatrix}$$

es la sustitución

$$A^{-1}=\begin{pmatrix}\alpha_1 & \alpha_2 & \dots & \alpha_n\\ 1 & 2 & \dots & n\end{pmatrix},$$

que se obtiene de la sustitución A, permutando la fila superior con la inferior.

Veamos ahora unas sustituciones de una forma especial, que se obtienen de la sustitución unidad E mediante una trasposición efectuada en su fila inferior. Tales sustituciones son impares, se llaman *trasposiciones* y tienen la forma:

$$\begin{pmatrix}\dots & i & \dots & j & \dots\\ \dots & j & \dots & i & \dots\end{pmatrix}, \tag{8}$$

donde los puntos suspensivos sustituyen a los símbolos que permanecen en su sitio. Convengamos en designar esta trasposición con la notación (i, j). La aplicación de la trasposición de los símbolos i, j a la fila inferior de la expresión (7) de una sustitución arbitraria A, es equivalente a multiplicar la sustitución A a la derecha por la sustitución (8), es decir, por (i, j). Ya sabemos que todas las permutaciones de n símbolos se pueden obtener de una de ellas, por ejemplo, de la permutación $1, 2, \dots, n$, realizando trasposiciones consecutivas; por eso, toda sustitución se puede obtener de la sustitución idéntica mediante la realización sucesiva de unas cuantas trasposiciones en la fila inferior, es decir, mediante una multiplicación sucesiva por sustituciones de la forma (6). Por consiguiente, se puede afirmar (omitiendo el factor E), que *toda sustitución se puede representar en forma de un producto de trasposiciones.*

Toda sustitución se puede descomponer de muchas maneras diversas en un producto de trasposiciones. Por ejemplo, siempre se pueden agregar dos factores iguales de la forma (i, j) (i, j) que, al multiplicarlos, darán las sustitución E, es decir, que se eliminan mutuamente. Señalemos un ejemplo menos trivial:

$$\begin{pmatrix}1 & 2 & 3 & 4 & 5\\ 2 & 5 & 4 & 3 & 1\end{pmatrix}=(12)\,(15)\,(34)=(14)\,(24)\,(45)\,(34)\,(13).$$

El nuevo método de determinación de la paridad de una sustitución se basa en el teorema siguiente:

En todas las descomposiciones de una sustitución en producto de trasposiciones, la paridad del número de estas trasposiciones es la misma, y coincide con la paridad de la sustitución misma.

Así, la sustitución del ejemplo considerado anteriormente es impar, como se puede comprobar calculando el número de inversiones.

El teorema quedará demostrado si se muestra que el *producto de cualesquiera k trasposiciones es una sustitución, cuya paridad*

coincide con la paridad del número k. Para $k = 1$ esto es cierto, puesto que una trasposición es una sustitución impar. Supongamos que ya está demostrada nuestra afirmación para el caso de $k - 1$ factores. Entonces, su validez para k factores se deducirá de que los números $k - 1$ y k son de paridad contraria, y el producto de una sustitución (en el caso considerado, del producto de los primeros $k - 1$ factores) por una trasposición es equivalente a la realización de esta trasposición en la fila inferior de la sustitución, es decir, cambia su paridad.

Un método muy cómodo de expresión de las sustituciones, que permite hallar fácilmente su paridad, es la *descomposición en ciclos*. Toda sustitución de grado n puede dejar en el sitio algunos de los símbolos $1, 2, \ldots, n$, otros, verdaderamente los puede transportar.

Una sustitución se llama *sustitución circular o ciclo* si al repetirla un número suficiente de veces, cada uno de los símbolos que verdaderamente se transportan puede ser transportado sobre cualquiera otro de estos símbolos. Tal es, por ejemplo, la sustitución de octavo grado

$$\begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \\ 1 & 8 & 6 & 4 & 5 & 2 & 7 & 3 \end{pmatrix};$$

ésta verdaderamente transporta los símbolos 2, 3, 6 y 8, a saber, el símbolo 2 sobre el 8, el símbolo 8 sobre el 3, el símbolo 3 sobre el 6 y el símbolo 6 de nuevo sobre el 2.

Todas las trasposiciones pertenecen al conjunto de los ciclos. Por analogía con la forma abreviada de expresión de las trasposiciones que se había empleado anteriormente, para los ciclos se usa la siguiente forma de expresión: los símbolos que verdaderamente son transportados se escriben entre paréntesis uno tras otro, en el mismo orden en que se sustituyen unos por otros al repetir la sustitución; la expresión comienza por cualquiera de los símbolos que verdaderamente se transportan y termina con el símbolo que se transporta sobre el primero. Así, para el ejemplo indicado anteriormente, esta expresión tiene la forma:

$$(2\ 8\ 3\ 6).$$

El número de símbolos que verdaderamente son transportados en el ciclo se llama *longitud* del mismo.

Se dice que dos ciclos de grado n son *independientes*, si no tienen símbolos comunes que verdaderamente sean transportados. Se comprende que, al multiplicar ciclos independientes, el orden de los factores no influye en el resultado.

Toda sustitución se puede descomponer de modo único en un producto de ciclos independientes dos a dos. La demostración de esta afirmación no representa dificultad alguna y la omitimos. La descomposición se realiza del modo siguiente: comenzamos por cualquiera de los símbolos que verdaderamente se transportan y escribimos tras él aquellos símbolos sobre los que éste se transporta al repetir la sustitución. Continuamos así, hasta que volvamos a obtener el símbolo inicial. Después de que «se cierre» este ciclo, comenzamos con uno de los símbolos que quedan y que verdaderamente se transportan, obteniendo así el segundo ciclo, etc.

Ejemplos

1) $\begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 3 & 5 & 1 & 2 & 4 \end{pmatrix} = (13)(254).$

2) $\begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \\ 5 & 2 & 8 & 7 & 6 & 1 & 4 & 3 \end{pmatrix} = (156)(38)(47).$

Recíprocamente, para cada sustitución, dada mediante su descomposición en ciclos independientes, se puede hallar una expresión en la forma ordinaria (con la condición de que se conozca el grado de la sustitución). Por ejemplo:

$$3)\ (1372)\,(45)=\begin{pmatrix}1&2&3&4&5&6&7\\3&1&7&5&4&6&2\end{pmatrix},$$

si se sabe que el grado de esta sustitución es igual a 7.

Sea dada una sustitución de grado n, y sea s el número de ciclos independientes en su descomposición, más el número de símbolos que permanecen en su sitio*. La diferencia n-s se llama *decremento* de la sustitución. Es evidente, que el decremento es igual al número de los símbolos que verdaderamente se transportan, menos el número de ciclos independientes que forman parte de la descomposición de la sustitución. Para los ejemplos 1), 2) y 3), considerados anteriormente, el decremento es igual a 3, 4, y 4, respectivamente.

La paridad de una sustitución coincide con la paridad del decremento de ella.

En efecto, todo ciclo de longitud k se puede representar en forma de un producto de $k-1$ trasposiciones del modo siguiente:

$$(i_1,\ i_2,\ \ldots,\ i_k)=(i_1,\ i_2)\,(i_1,\ i_3)\ \ldots\ (i_1,\ i_k).$$

Supongamos dada la descomposición de la sustitución A en ciclos independientes. Si se descompone cada uno de los ciclos en el producto de las trasposiciones que acabamos de indicar, obtendremos la expresión de la sustitución A en forma de un producto de trasposiciones. El número de estas trasposiciones será, evidentemente, menor que el número de los símbolos que verdaderamente son transportados por la sustitución A, en un número igual al número de los ciclos independientes en la descomposición de la sustitución. De aquí se deduce, que la sustitución A se puede descomponer en un producto de trasposiciones, cuyo número es igual al decremento. Por consiguiente, la paridad de la sustitución se determina por la paridad del decremento.

§ 4. Determinantes de n-ésimo orden

Queremos generalizar ahora para el caso de un n arbitrario, los resultados obtenidos en el § 2 para $n = 2$ y 3. Con este fin, es necesario definir los determinantes de n-ésimo orden. Sin embargo, es imposible hacer esto del mismo modo que se introdujeron los determinantes de segundo y tercer orden, es decir, resolviendo en forma general un sistema de ecuaciones lineales, pues, a medida que aumentase n, los cálculos se harían más y más complicados, y siendo n arbitrario, éstos serían prácticamente irrealizables. Procederemos de otro modo. Examinaremos los determinantes de segundo y tercer orden ya conocidos. Procuraremos establecer una ley general, de acuerdo a la cual se expresan estos determinantes mediante los elementos de las matrices correspondientes y tomaremos esta ley por definición para el determinante de orden n. Después demostraremos que con esta definición sigue cumpliéndose la regla de Cramer.

---

* A todo símbolo que se mantiene en su sitio se podía haber puesto en correspondencia un «ciclo» de longitud 1, es decir, que en el ejemplo 2), indicado anteriormente, se podría escribir: (156) (38) (47) (2). Sin embargo, no procederemos de este modo.

Recordemos las expresiones de los determinantes de segundo y tercer orden:

$$\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11}a_{22} - a_{12}a_{21},$$

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} - \\ - a_{13}a_{22}a_{31} - a_{12}a_{21}a_{33} - a_{11}a_{23}a_{32}.$$

Obsérvese que todo término del determinante de segundo orden es un producto de dos elementos, situados en diversas filas y en diversas columnas. Además, todos los productos de este tipo que se pueden formar con los elementos de la matriz de segundo orden (en total son dos), se han utilizado como términos del determinante. De modo semejante, todo término del determinante de tercer orden representa un producto de tres elementos, tomados también uno a uno de cada fila y de cada columna. Todos los productos de estos se utilizan también como términos del determinante.

Sea dada ahora una matriz cuadrada de orden n

$$\begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{pmatrix}. \tag{1}$$

Consideremos todos los productos posibles de n elementos de esta matriz, situados en diferentes filas y en diferentes columnas, o sea, los productos de la forma

$$a_{1\alpha_1}\, a_{2\alpha_2} \dots a_{n\alpha_n}, \tag{2}$$

donde los subíndices $\alpha_1, \alpha_2, \dots, \alpha_n$ forman una de las permutaciones de los números $1, 2, \dots, n$. El número de estos productos es igual al número de las diversas permutaciones de n símbolos, es decir, es igual a $n!$. Vamos a tomar todos estos productos por términos del futuro determinante de n-ésimo orden, correspondiente a la matriz (1).

Para determinar el signo con que figura el producto (2) en el determinante, observemos que con los subíndices de este producto se puede formar la sustitución

$$\begin{pmatrix} 1 & 2 & \dots & n \\ \alpha_1 & \alpha_2 & \dots & \alpha_n \end{pmatrix}, \tag{3}$$

donde i se sustituye por α_i, si el elemento situado en la i-ésima fila y en la α_i-ésima columna de la matriz (1) forma parte del producto (2). Examinando las expresiones de los determinantes de segundo y tercer orden, observamos que en ellos figuran con signo más los

términos cuyos subíndices forman una sustitución par, y con signo menos, los términos cuyos subíndices forman una sustitución impar. Resulta natural conservar también esta ley en la definición del determinante de n-ésimo orden.

Por lo tanto, llegamos a la siguiente definición: se llama determinante de n-ésimo orden, correspondiente a la matriz (1), a la suma algebraica de $n!$ términos, constituida del modo siguiente: son términos de ella todos los productos posibles de n elementos de la matriz, tomados uno de cada fila y de cada columna, tomando el término *con signo más, si sus subíndices forman una sustitución par, y con* signo menos, en el caso contrario.

Para escribir el determinante de n-ésimo orden correspondiente a la matriz (1) se empleará la notación que se usó en el caso de los determinantes de segundo y tercer orden:

$$\begin{vmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \cdot & \cdot & \cdot & \cdot \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{vmatrix}. \tag{4}$$

Los determinantes de n-ésimo orden, para $n = 2$ y $n = 3$, se convierten en los determinantes de segundo y tercer orden considerados anteriormente; para $n = 1$, es decir, para las matrices constituidas de un sólo elemento, el determinante es igual al elemento mismo. Sin embargo, por ahora, todavía no sabemos si para $n > 3$ se pueden utilizar los determinantes de n-ésimo orden para la resolución de sistemas de ecuaciones lineales. Esto se mostrará en el § 7; pero previamente tenemos que estudiar detalladamente los determinantes de n-ésimo orden y, en particular, tenemos que hallar un método para su cálculo, puesto que sería muy difícil calcular los determinantes partiendo de su definición, incluso para n no muy grandes.

Ahora estableceremos las propiedades elementales de los determinantes de n-ésimo orden, relativas fundamentalmente a una de las dos cuestiones. Por una parte, nos interesarán las condiciones para que el determinante sea igual a cero; por otra parte, señalaremos unas transformaciones de la matriz que no alteran a su determinante o que proporcionan una alteración de éste, fácilmente calculable.

Llamaremos *transposición* de la matriz A a una transformación de la misma, según la cual sus filas se sustituyen por sus columnas del mismo orden, es decir, el paso de la matriz (1) a la matriz

$$\begin{pmatrix} a_{11} & a_{21} & \dots & a_{n1} \\ a_{12} & a_{22} & \dots & a_{n2} \\ \cdot & \cdot & \cdot & \cdot \\ a_{1n} & a_{2n} & \dots & a_{nn} \end{pmatrix}; \tag{5}$$

se puede decir que transponer la matriz (1) es hacerla girar alrededor de la diagonal principal. Correspondientemente, se dice, que el determinante

$$\begin{vmatrix} a_{11} & a_{21} & \dots & a_{n1} \\ a_{12} & a_{22} & \dots & a_{n2} \\ \cdot & \cdot & \cdot & \cdot \\ a_{1n} & a_{2n} & \dots & a_{nn} \end{vmatrix} \tag{6}$$

se obtiene transponiendo el determinante (4).

Propiedad 1. *El determinante no varía al transponerlo.*

En efecto, todo término del determinante (4) es de la forma

$$a_{1\alpha_1} a_{2\alpha_2} \dots a_{n\alpha_n}, \tag{7}$$

donde los segundos subíndices forman una permutación de los símbolos 1, 2, ..., n. Pero, todos los factores del producto (7) se mantienen también en el determinante (6) en diferentes filas y en diferentes columnas, es decir, que (7) es también un término del determinante transpuesto. Es evidente que lo recíproco también es justo. Por lo tanto, los determinantes (4) y (6) están constituidos por los mismos términos. El signo del término (7) en el determinante (4) se determina por la paridad de la sustitución

$$\begin{pmatrix} 1 & 2 & \dots & n \\ \alpha_1 & \alpha_2 & \dots & \alpha_n \end{pmatrix}; \tag{8}$$

en el determinante (6), los primeros subíndices de los elementos indican el número de orden de la columna, mientras que los segundos subíndices indican el número de orden de la fila. Por consiguiente, en el determinante (6) al término (7) corresponde la sustitución

$$\begin{pmatrix} \alpha_1 & \alpha_2 & \dots & \alpha_n \\ 1 & 2 & \dots & n \end{pmatrix}. \tag{9}$$

Por lo general, las sustituciones (8) y (9) son diferentes, pero, evidentemente, tienen una misma paridad y, por lo tanto, el término (7) tiene un mismo signo en ambos determinantes. Por consiguiente, los determinantes (4) y (6) representan sumas de términos iguales, tomados con signos iguales, es decir, son iguales entre sí.

De la propiedad 1 se deduce que cualquier afirmación sobre las filas del determinante es válida también para sus columnas y viceversa, es decir, en el *determinante* (a distinción de las matrices), las *filas y las columnas gozan de los mismos derechos.* Partiendo de esto, las siguientes ocho propiedades (2-9) se enunciarán y se demostrarán solamente para las filas del determinante; las propiedades análogas para las columnas no necesitarán una demostración especial.

Propiedad 2. *Si una de las filas del determinante está constituida por ceros, el determinante es igual a cero.*

En efecto, *supongamos que* todos los elementos de la *i*-ésima fila del determinante son iguales a cero. En cada uno de los términos del determinante tiene que estar incluido uno de los elementos de la *i*-ésima fila, por lo cual, en nuestro caso, todos los términos del determinante son iguales a cero.

Propiedad 3. *Si un determinante se obtiene de otro permutando dos filas, todos los términos del primer determinante serán términos del segundo, pero con signos contrarios, es decir, al permutar dos filas, el determinante sólo cambia de signo.*

En efecto, supongamos que en el determinante (4) se permutan la *i*-ésima y la *j*-ésima filas, $i \neq j$, y que todas las demás filas se mantienen en su sitio. Obtenemos el determinante

$$\begin{vmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ \cdot & \cdot & \cdot & \cdot \\ a_{j1} & a_{j2} & \dots & a_{jn} \\ \cdot & \cdot & \cdot & \cdot \\ a_{i1} & a_{i2} & \dots & a_{in} \\ \cdot & \cdot & \cdot & \cdot \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{vmatrix} \begin{matrix} \\ \\ (i) \\ \\ (j) \\ \\ \\ \end{matrix} \qquad (10)$$

(al margen están señalados los números de las filas). Si

$$a_{1\alpha_1} a_{2\alpha_2} \dots a_{n\alpha_n} \qquad (11)$$

es un término del determinante (4), evidentemente, todos sus factores se mantienen también en el determinante (10) en diferentes filas y columnas. Por lo tanto, los determinantes (4) y (10) constan de los mismos términos. En el determinante (4) al término (11) le corresponde la sustitución

$$\begin{pmatrix} 1 & 2 & \dots & i & \dots & j & \dots & n \\ \alpha_1 & \alpha_2 & \dots & \alpha_i & \dots & \alpha_j & \dots & \alpha_n \end{pmatrix}, \qquad (12)$$

mientras que en el determinante (10), la sustitución

$$\begin{pmatrix} 1 & 2 & \dots & j & \dots & i & \dots & n \\ \alpha_1 & \alpha_2 & \dots & \alpha_i & \dots & \alpha_j & \dots & \alpha_n \end{pmatrix}, \qquad (13)$$

puesto que el elemento $a_{i\alpha_i}$, por ejemplo, está ahora en la *j*-ésima fila, pero se mantiene en la α_i*-ésima columna anterior. Sin embargo,* la sustitución (13) se obtiene de la sustitución (12) mediante una trasposición en la fila superior, o sea, tiene paridad contraria. De esto se deduce, que todos los términos del determinante (4) forman parte del determinante (10), pero con signos contrarios, es decir, los determinantes (4) y (10) se diferencian entre sí solamente en el signo.

Propiedad 4. *Un determinante que tiene dos filas iguales es igual a cero.*

En efecto, supongamos que el valor del determinante es igual a d y que son iguales entre sí los elementos correspondientes de su i-ésima y j-ésima filas ($i \neq j$). En virtud de la propiedad 3, después de permutar estas dos filas, el determinante se hace igual a $-d$. Sin embargo, como las filas que se permutan son iguales el determinante, en realidad, no varía, o sea, $d = -d$; de donde $d = 0$.

Propiedad 5. *Si se multiplican todos los elementos de una fila del determinante por un número k, el mismo determinante queda multiplicado por k.*

Supongamos que se han multiplicado por k todos los elementos de la i-ésima fila. Cada término del determinante contiene exactamente un elemento de la i-ésima fila. Por lo tanto, todo términc adquiere el factor k, es decir, el mismo determinante queda multiplicado por k.

Esta propiedad también se puede expresar así: *el factor común de todos los elementos de una fila del determinante se puede sacar fuera del signo de éste.*

Propiedad 6. *Un determinante que tiene dos filas proporcionales es igual a cero.*

Supongamos que los elementos de la j-ésima fila del determinante se diferencian de los elementos correspondientes de la i-ésima fila ($i \neq j$) en un mismo factor k. Sacando este factor común k de la j-ésima fila fuera del signo del determinante, obtenemos un determinante con dos filas iguales. Este será igual a cero, por la propiedad 4.

La propiedad 4, así como la propiedad 2 para $n > 1$, son, evidentemente, casos particulares de la propiedad 6 (para $k = 1$ y $k = 0$).

Propiedad 7. *Si todos los elementos de la i-ésima fila de un determinante de n-ésimo orden representan una suma de dos sumandos:*

$$a_{ij} = b_j + c_j, \quad j = 1, \ldots, n,$$

el determinante es igual a la suma de dos determinantes, en los que todas las filas, menos la i-ésima, coinciden con las del determinante dado, mientras que la i-ésima fila de uno de los sumandos consta de los elementos b_j y la del otro, de los elementos c_j.

Todo término del determinante dado se puede representar de la forma

$$a_{1\alpha_1}a_{2\alpha_2}\ldots a_{i\alpha_i}\ldots a_{n\alpha_n} = a_{1\alpha_1}a_{2\alpha_2}\ldots (b_{\alpha_i}+c_{\alpha_i})\ldots a_{n\alpha_n} =$$
$$= a_{1\alpha_1}a_{2\alpha_2}\ldots b_{\alpha_i}\ldots a_{n\alpha_n} + a_{1\alpha_1}a_{2\alpha_2}\ldots c_{\alpha_i}\ldots a_{n\alpha_n}.$$

Rouniendo los primeros términos de estas sumas (con los mismos signos que tenían los términos correspondientes en el determinante

dado), obtenemos un determinante de orden n, que solamente se diferencia del dado, en que en la i-ésima fila, en lugar de los elementos a_{ij}, figuran los elementos b_j. Correspondientemente, los segundos sumandos forman un determinante en cuya i-ésima fila figuran los elementos c_j. Por lo tanto,

$$\begin{vmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ \dots & \dots & \dots & \dots \\ b_1+c_1 & b_2+c_2 & \dots & b_n+c_n \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{vmatrix} = \begin{vmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ \dots & \dots & \dots & \dots \\ b_1 & b_2 & \dots & b_n \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{vmatrix} + \begin{vmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ \dots & \dots & \dots & \dots \\ c_1 & c_2 & \dots & c_n \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{vmatrix}.$$

La propiedad 7 se generaliza sin dificultad al caso en que todo elemento de la i-ésima fila es una suma, no de dos, sino de m sumandos, $m \geqslant 2$.

Se dice que la i-ésima fila de un determinante es *combinación lineal* de las demás filas, si para cada fila del número de orden j, $j = 1, \dots, i-1, i+1, \dots, n$, se puede señalar un número k_j tal, que multiplicando la j-ésima fila por k_j y agregando después todas las filas, menos la i-ésima (la suma de las filas se debe entender como la suma por separado de los elementos de todas estas filas en cada columna), se obtiene la i-ésima fila. Algunos de los coeficientes k_j pueden ser iguales a cero, es decir, en realidad, la i-ésima fila es combinación lineal, no de todas, sino de algunas filas restantes. En particular, si solamente uno de los coeficientes k_j es diferente de cero, obtenemos el caso de proporcionalidad de dos filas. Finalmente, si una fila se compone totalmente de ceros, ésta siempre será combinación lineal de las demás filas: caso en que todos los k_j son iguales a cero.

Propiedad 8. *Si una de las filas del determinante es combinación lineal de las demás, el determinante es igual a cero.*

Sea, por ejemplo, la i-ésima fila, combinación lineal de las otras s filas, $1 \leqslant s \leqslant n-1$. Entonces, todo elemento de la i-ésima fila será una suma de s términos. Por lo tanto, aplicando la propiedad 7, representamos nuestro determinante en forma de una suma de determinantes, en cada uno de los cuales la i-ésima fila será proporcional a una de las otras filas. Según la propiedad 6, todos estos determinantes son iguales a cero; por consiguiente, también será igual a cero el determinante dado.

Esta propiedad es una generalización de la propiedad 6, y, como se demostrará en el § 10, es el caso más general de igualdad a cero del determinante.

Propiedad 9. *El determinante no varía si a los elementos de una de sus filas se agregan los elementos correspondientes de otra fila, multiplicados por un mismo número.*

Supongamos que a la i-ésima fila del determinante d se le agrega la j-ésima fila, $j \neq i$, multiplicada por el número k, es decir, que en el nuevo determinante todo elemento de la i-ésima fila tiene la forma $a_{is} + ka_{js}$, $s = 1, 2, \ldots, n$. Entonces, de acuerdo a la propiedad 7, este determinante es igual a la suma de dos determinantes, el primero de los cuales es d, mientras que el segundo contiene dos filas proporcionales y, por ello, es igual a cero.

Como el número k puede ser negativo, el *determinante tampoco variará al restar de una de sus filas otra fila, multiplicada por un número*. En general, *el determinante no varía si a una de sus filas se agrega cualquier combinación lineal de las demás.*

Veamos el siguiente ejemplo. Un determinante se llama *antisimétrico*, si sus elementos, situados simétricamente respecto de la diagonal principal, se diferencian entre sí solamente en el signo, es decir, si para todos i y j se tiene $a_{ji} = -a_{ij}$; de esto se deduce que para todo i será $a_{ii} = -a_{ii} = 0$. Por lo tanto, el determinante tiene la forma

$$d = \begin{vmatrix} 0 & a_{12} & a_{13} & \ldots & a_{1n} \\ -a_{12} & 0 & a_{23} & \ldots & a_{2n} \\ -a_{13} & -a_{23} & 0 & \ldots & a_{3n} \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ -a_{1n} & -a_{2n} & -a_{3n} & \ldots & 0 \end{vmatrix}.$$

Multiplicando cada fila de este determinante por -1, obtenemos el determinante transpuesto, que es de nuevo igual a d, de donde, en virtud de la propiedad 5, resulta:

$$(-1)^n d = d.$$

Para n impar, se deduce que: $-d = d$, es decir, $d = 0$. Por lo tanto, *todo determinante antisimétrico (o hemisimétrico) de orden impar es igual a cero.*

§ 5. Los menores y sus complementos algebraicos

Antes se había indicado que sería difícil calcular un determinante de n-ésimo grado aplicando directamente su definición, o sea, escribiendo cada vez todos los $n!$ términos, determinando sus signos, etc. Existen métodos más sencillos para calcular los determinantes, basados en el hecho de que un determinante de orden n se puede expresar mediante determinantes de órdenes inferiores. Introduzcamos, con este fin, el siguiente concepto.

Sea dado un determinante d de orden n. Tomemos un número entero k que satisfaga la condición $1 \leqslant k \leqslant n - 1$, y elijamos arbitrariamente en el determinante d, k filas y k columnas. Los elementos situados en las intersecciones de estas filas y de estas columnas, es decir, pertenecientes a una de las filas y a una de las columnas elegidas, forman, evidentemente, una matriz de orden k. El determinante de esta matriz se llama *menor de orden k* del determinante d. Se puede decir también que el menor de orden k es el determinante

que se obtiene después de suprimir $n-k$ filas y $n-k$ columnas en el determinante d. En particular, después de haber suprimido en el determinante una fila y una columna, obtenemos un menor de orden $(n-1)$; por otra parte, los mismos elementos del determinante d por separado representan menores de primer orden.

Supongamos que en un determinante d de n-ésimo orden se ha tomado un menor M de orden k. Suprimiendo las filas y columnas, en cuyas intersecciones figura este menor, resulta un menor M' de $(n-k)$-ésimo orden, denominado *menor complementario* del menor M. Suprimiendo, por el contrario, las filas y columnas en las que están situados los elementos del menor M', obtendremos el menor M. Por lo tanto, se puede hablar de un *par de menores complementarios entre sí* del determinante. En particular, el elemento a_{ij} y el menor de $(n-1)$-ésimo orden que se obtiene suprimiendo en el determinante la i-ésima fila y la j-ésima columna, formarán un par de menores complementarios entre sí.

Si un menor M de k-ésimo orden está situado en las filas de orden $i_1, i_2, \ldots, i_k$ y en las columnas de orden $j_1, j_2, \ldots, j_k$, entonces, denominaremos *complemento algebraico* del menor M a su menor complementario M', tomado con el signo más o menos, según que sea par o impar la suma de los números de orden de todas las filas y columnas en las que está situado el menor M, es decir la suma

$$s_M = i_1 + i_2 + \ldots + i_k + j_1 + j_2 + \ldots j_k. \qquad (1)$$

En otras palabras, el complemento algebraico del menor M es el número $(-1)^{s_M} M'$.

El producto de cualquier menor M de k-ésimo orden por su complemento algebraico en el determinante d es una suma algebraica, cuyos sumandos, obtenidos al multiplicar los términos del menor M por los términos del menor complementario M' tomados con el signo $(-1)^{s_M}$, son ciertos términos del determinante d, coincidiendo sus signos en esta suma con los signos que tienen en el determinante.

Comenzaremos la demostración de este teorema con el caso en que el menor M está situado en el ángulo superior de la izquierda del determinante:

$$d = \left|\begin{array}{ccc|ccc} a_{11} & \ldots & a_{1k} & a_{1,k+1} & \ldots & a_{1n} \\ \ldots & M & \ldots & \ldots & \ldots & \ldots \\ a_{k1} & \ldots & a_{kk} & a_{k,k+1} & \ldots & a_{kn} \\ \hline a_{k+1,1} & \ldots & a_{k+1,k} & a_{k+1,k+1} & \ldots & a_{k+1,n} \\ \ldots & \ldots & \ldots & \ldots & M' & \ldots \\ a_{n1} & \ldots & a_{nk} & a_{n,k+1} & \ldots & a_{nn} \end{array}\right|,$$

es decir, en las filas cuyos números de orden son $1, 2, \ldots, k$ y en las columnas que tienen los mismos números de orden. Entonces,

el menor M' ocupará el ángulo inferior de la derecha del determinante. En este caso, el número s_M es par:

$$s_M = 1 + 2 + \ldots + k + 1 + 2 + \ldots + k = 2(1 + 2 + \ldots + k),$$

por eso, el mismo menor M' sirve de complemento algebraico para M.

Tomemos un término arbitrario del menor M

$$a_{1\alpha_1} a_{2\alpha_2} \ldots a_{k\alpha_k}; \tag{2}$$

su signo en M será $(-1)^l$, donde l es el número de inversiones en la sustitución

$$\begin{pmatrix} 1 & 2 & \ldots & k \\ \alpha_1 & \alpha_2 & \ldots & \alpha_k \end{pmatrix}. \tag{3}$$

El término arbitrario del menor M'

$$a_{k+1,\ \beta_{k+1}}\ a_{k+2,\ \beta_{k+2}} \ldots a_{n\beta_n} \tag{4}$$

tiene en éste el signo $(-1)^{l'}$, donde l' es el número de inversiones en la sustitución

$$\begin{pmatrix} k+1 & k+2 & \ldots & n \\ \beta_{k+1} & \beta_{k+2} & \ldots & \beta_n \end{pmatrix}. \tag{5}$$

Multiplicando los términos (2) y (4), obtenemos el producto de n elementos

$$a_{1\alpha_1} a_{2\alpha_2} \ldots a_{k\alpha_k} a_{k+1,\ \beta_{k+1}} a_{k+2,\ \beta_{k+2}} \ldots a_{n\beta_n}, \tag{6}$$

situados en diferentes filas y columnas del determinante; por consiguiente, éste será un término del determinante d. El signo del término (6) en el producto MM' será igual al producto de los signos de los términos (2) y (4), o sea, $(-1)^l \cdot (-1)^{l'} = (-1)^{l+l'}$. Sin embargo, el término (6) tiene también este mismo signo en el determinante d. En efecto, la fila inferior de la sustitución

$$\begin{pmatrix} 1 & 2 & \ldots & k & k+1 & k+2 & \ldots & n \\ \alpha_1 & \alpha_2 & \ldots & \alpha_k & \beta_{k+1} & \beta_{k+2} & \ldots & \beta_n \end{pmatrix},$$

formada por los índices de este término, contiene solamente $l + l'$ inversiones, puesto que ningún α puede formar inversión con ningún β: todos los α no son mayores que k, mientras que todos los β no son menores que $k + 1$.

De este modo, queda demostrado el caso particular considerado del teorema. Pasemos a examinar el caso general. Supongamos que el menor M está situado en las filas que tienen los números de orden

$i_1, i_2, \ldots, i_k$ y en las columnas que tienen los números de orden $j_1, j_2, \ldots, j_k$, siendo

$$i_1 < i_2 < \ldots < i_k, \quad j_1 < j_2 < \ldots < j_k.$$

Trasponiendo las filas y las columnas, procuremos llevar el menor M al ángulo superior de la izquierda, de modo que no se altere el menor complementario. Con este fin, trasponemos la i_1-ésima fila con la (i_1-1)-ésima, después, con la (i_1-2)-ésima, etc., hasta que la i_1-ésima fila ocupe el lugar de la primera; para esto, tendremos que trasponer las filas i_1-1 veces. Después, trasponemos sucesivamente la i_2-ésima fila con todas las filas situadas sobre ella, hasta que se sitúe directamente debajo de la i_1-ésima fila, es decir, en el sitio que ocupaba la segunda fila antes de todas las transformaciones; como es fácil comprobar, para ello tenemos que trasponer las filas i_2-2 veces. De modo análogo, trasladamos la i_3-ésima fila al lugar de la tercera fila, etc., hasta que la i_k-ésima ocupe el lugar de la k-ésima fila. En total, tendremos que efectuar

$$(i_1-1)+(i_2-2)+\ldots+(i_k-k)=$$
$$=(i_1+i_2+\ldots+i_k)-(1+2+\ldots+k)$$

trasposiciones de las filas.

El menor M ya está situado en las primeras k filas del nuevo determinante. Ahora trasponemos sucesivamente las columnas del determinante: la j_1-ésima con todas las precedentes hasta que ocupe el primer lugar, después, la j_2-ésima, hasta que ocupe el segundo lugar, etc. En total, las columnas serán traspuestas

$$(j_1+j_2+\ldots+j_k)-(1+2+\ldots+k)$$

veces.

Después de todas estas transformaciones llegamos a un determinante nuevo d', en el cual, el menor M ocupa el ángulo superior de la izquierda. Como habíamos traspuesto cada vez solamente las filas y columnas contiguas, no sufrirá ninguna alteración la colocación mutua de las filas y columnas que contenían el menor M' en el determinante d. Por lo tanto, el menor M' se mantiene también como menor complementario del menor M en el determinante d', ocupando ya, sin embargo, el ángulo inferior de la derecha. Como hemos demostrado, el producto MM' es una suma de cierto número de términos del determinante d', tomados con los mismos signos que tenían en d'. No obstante, el determinante d' se ha obtenido del determinante d mediante

$$[(i_1+i_2+\ldots+i_k)-(1+2+\ldots+k)]+$$
$$+[(j_1+j_2+\ldots+j_k)-(1+2+\ldots+k)]=s_M-2(1+2+\ldots+k)$$

trasposiciones de las filas y columnas. Por ello, como sabemos por el párrafo anterior, los términos del determinante d' solamente se diferencian de los términos correspondientes del determinante d en el signo $(-1)^{s_M}$ (se comprende que el número par $2(1+2+\ldots+k)$ no influye en el signo). De aquí se deduce que el producto $(-1)^{s_M}MM'$ se compone de una cierta cantidad de términos del determinante d, tomados con los mismos signos que tenían en este determinante. De esta manera, el teorema queda demostrado.

Obsérvese que si los menores M y M' son complementarios entre sí, los números s_M y $s_{M'}$ son de una misma paridad. En efecto, el número de orden de cada fila y de cada columna está incluido como sumando en uno, y sólo en uno, de estos números. Por consiguiente, la suma $s_M + s_{M'}$ es igual a la suma de los números de orden de todas las filas y columnas del determinante, es decir, es igual a la paridad del número $2(1+2+\ldots+n)$.

§ 6. Cálculo de determinantes

Los resultados del párrafo anterior ofrecen la posibilidad de reducir el cálculo de un determinante de n-ésimo orden al cálculo de unos cuantos determinantes de $(n-1)$-ésimo orden. Introduzcamos, primero, las siguientes notaciones: si a_{ij} es un elemento del determinante d, designaremos con M_{ij} el menor complementario, o abreviando, el *menor de este elemento*, es decir, el menor de $(n-1)$-ésimo orden obtenido después de suprimir la i-ésima fila y la j-ésima columna en el determinante. Designaremos con A_{ij} el complemento algebraico del elemento a_{ij},

$$A_{ij} = (-1)^{i+j} M_{ij}.$$

Como se ha demostrado anteriormente, el producto $a_{ij}A_{ij}$ representa una suma de unos cuantos términos del determinante d, incluidos en esta suma con los mismos signos que tenían en el determinante d. Es fácil calcular el número de estos términos: es igual al número de términos en el menor M_{ij}, es decir, es igual a $(n-1)!$

Elijamos ahora una fila i-ésima cualquiera del determinante d y tomemos el producto de cada elemento de esta fila por su complemento algebraico:

$$a_{i1}A_{i1}, \quad a_{i2}A_{i2}, \ \ldots, \ a_{in}A_{in}. \tag{1}$$

Ningún término del determinante d puede estar incluido en dos productos diferentes (1): todos los términos del determinante incluidos en el producto $a_{i1}A_{i1}$ contienen el elemento a_{i1} de la i-ésima fila.

Por ello, se diferencian de los términos que forman parte del producto $a_{i2}A_{i2}$, que contienen el elemento a_{i2} de la *i*-ésima fila, etc.

Por otra parte, el número total de términos del determinante d, incluidos en todos los productos (1), es igual a

$$(n-1)!\cdot n=n!.$$

Con éstos se agotan por completo todos los términos del determinante d. Por lo tanto, hemos demostrado que se verifica el siguiente *desarrollo del determinante d por los elementos de la i-ésima fila:*

$$d_i=a_{i1}A_{i1}+a_{i2}A_{i2}+\ldots a_{in}A_{in}, \tag{2}$$

lo que significa que el *determinante d es igual a la suma de los productos de todos los elementos de una fila arbitraria de él por sus complementos algebraicos.* Se puede obtener un desarrollo análogo del determinante por los elementos de cualquiera de sus columnas.

Sustituyendo en el desarrollo (2) los complementos algebraicos por los menores correspondientes con los signos más o menos, **reduciremos el cálculo del determinante de n-ésimo orden al cálculo de unos cuantos determinantes de (n-1)-ésimo orden.** Obsérvese que si algunos de los elementos de la i-ésima fila son iguales a cero, no habrá que calcular, naturalmente, sus menores correspondientes. En virtud de esto, es conveniente transformar previamente el determinante, aplicando la propiedad 9 (véase el § 4), para que en una de las filas o de las columnas haya un número suficientemente grande de elementos sustituidos por ceros. En realidad, la *propiedad 9 da la posibilidad de sustituir por ceros todos los elementos, menos uno, de cualquier fila o de cualquier columna.* En efecto, si $a_{ik}\neq 0$, cualquier elemento a_{ij}, $j\neq k$, de la i-ésima fila quedará sustituido por cero después de restar de la j-ésima columna la k-ésima columna multiplicada por $\frac{a_{ij}}{a_{ik}}$. De este modo, el cálculo de un determinante de n-ésimo orden se puede reducir al cálculo de **un solo** determinante de $(n-1)$-ésimo orden.

Ejemplos.

1. Calcular el determinante de cuarto orden

$$d=\begin{vmatrix} 3 & 1 & -1 & 2\\ -5 & 1 & 3 & -4\\ 2 & 0 & 1 & -1\\ 1 & -5 & 3 & -3\end{vmatrix}.$$

Desarrollémoslo por los elementos de la tercera fila, aprovechando la existencia de un cero:

$$d=(-1)^{3+1}\cdot 2\cdot\begin{vmatrix} 1 & -1 & 2 \\ 1 & 3 & -4 \\ -5 & 3 & -3 \end{vmatrix}+$$

$$+(-1)^{3+3}\cdot 1\cdot\begin{vmatrix} 3 & 1 & 2 \\ -5 & 1 & -4 \\ 1 & -5 & -3 \end{vmatrix}+(-1)^{3+4}\cdot(-1)\cdot\begin{vmatrix} 3 & 1 & -1 \\ -5 & 1 & 3 \\ 1 & -5 & 3 \end{vmatrix}.$$

Calculando los determinantes obtenidos de tercer orden, obtenemos:

$$d=2\cdot 16-40+48=40.$$

2. Calcular el determinante de quinto orden

$$d=\begin{vmatrix} -2 & 5 & 0 & -1 & 3 \\ 1 & 0 & 3 & 7 & -2 \\ 3 & -1 & 0 & 5 & -5 \\ 2 & 6 & -4 & 1 & 2 \\ 0 & -3 & -1 & 2 & 3 \end{vmatrix}.$$

Agregando a la *segunda fila la quinta*, ***multiplicada*** *por tres*, y restando de la cuarta fila la quinta, multiplicada por cuatro, obtenemos:

$$d=\begin{vmatrix} -2 & 5 & 0 & -1 & 3 \\ 1 & -9 & 0 & 13 & 7 \\ 3 & -1 & 0 & 5 & -5 \\ 2 & 18 & 0 & -7 & -10 \\ 0 & -3 & -1 & 2 & 3 \end{vmatrix}.$$

Desarrollando este determinante por los elementos de la tercera columna, que contiene solamente un elemento diferente de cero (con la suma de índices $5+3$, *es decir, par), obtenemos:*

$$d=(-1)\cdot\begin{vmatrix} -2 & 5 & -1 & 3 \\ 1 & -9 & 13 & 7 \\ 3 & -1 & 5 & -5 \\ 2 & 18 & -7 & -10 \end{vmatrix}.$$

Transformamos de nuevo el determinante obtenido, agregando a la primera fila la segunda, multiplicada por dos, restando de la tercera fila la segunda, multiplicada por tres, y de la cuarta, la segunda multiplicada por dos:

$$d=-\begin{vmatrix} 0 & -13 & 25 & 17 \\ 1 & -9 & 13 & 7 \\ 0 & 26 & -34 & -26 \\ 0 & 36 & -33 & -24 \end{vmatrix}.$$

Después, desarrollamos éste por los elementos de la primera columna, teniendo además en cuenta, que al único elemento de esta columna, diferente de cero, le corresponde una suma impar de índices. Resulta:

$$d=\begin{vmatrix} -13 & 25 & 17 \\ 26 & -34 & -26 \\ 36 & -33 & -24 \end{vmatrix}.$$

Calculemos este determinante de tercer orden, desarrollándolo previamente por los elementos de su tercera fila:

$$d=36\cdot\begin{vmatrix} 25 & 17 \\ -34 & -26 \end{vmatrix}-(-33)\cdot\begin{vmatrix} -13 & 17 \\ 26 & -26 \end{vmatrix}+(-24)\cdot\begin{vmatrix} -13 & 25 \\ 26 & -34 \end{vmatrix}=$$
$$=36\cdot(-72)-(-33)\cdot(-104)+(-24)\cdot(-208)=-1032.$$

3. *Si todos los elementos de un determinante, situados a un lado de la diagonal principal, son iguales a cero, el determinante es igual al producto de los elementos situados en la diagonal principal.*

Para un determinante de segundo orden, esta afirmación es evidente. Por ello, la vamos a demostrar por el método de inducción; supongamos que está demostrada ya para los determinantes de $(n$-$1)$-ésimo orden. Consideremos el determinante de n-ésimo orden:

$$d=\begin{vmatrix} a_{11} & a_{12} & a_{13} & \dots & a_{1n} \\ 0 & a_{22} & a_{23} & \dots & a_{2n} \\ 0 & 0 & a_{33} & \dots & a_{3n} \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & 0 & \dots & a_{nn} \end{vmatrix}.$$

Desarrollándolo por los elementos de la primera columna, obtenemos

$$d=a_{11}\cdot\begin{vmatrix} a_{22} & a_{23} & \dots & a_{2n} \\ 0 & a_{33} & \dots & a_{3n} \\ \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & \dots & a_{nn} \end{vmatrix}.$$

Al menor que figura en el segundo miembro se le puede aplicar la hipótesis de inducción, es decir, es igual a $a_{22}\,a_{33}\dots a_{nn}$; de donde

$$d=a_{11}a_{22}a_{33}\dots a_{nn}.$$

4. Se llama *determinante de Vandermonde* al siguiente:

$$d=\begin{vmatrix} 1 & 1 & 1 & \dots & 1 \\ a_1 & a_2 & a_3 & \dots & a_n \\ a_1^2 & a_2^2 & a_3^2 & \dots & a_n^2 \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ a_1^{n-1} & a_2^{n-1} & a_3^{n-1} & \dots & a_n^{n-1} \end{vmatrix}.$$

Demostremos que para *cualquier* n *el determinante de Vandermonde es igual al producto de todas las diferencias posibles* $a_i - a_j$, donde $1 \leqslant j < i \leqslant n$. En efecto, para $n = 2$, se tiene

$$\begin{vmatrix} 1 & 1 \\ a_1 & a_2 \end{vmatrix} = a_2 - a_1.$$

Supongamos que nuestra afirmación está demostrada ya para los determinantes de Vandermonde de $(n-1)$-ésimo orden. Transformemos el determinante d del modo siguiente: de la n-ésima (la última) fila restamos la $(n-1)$-ésima, multiplicada por a_1; después de la $(n-1)$-ésima restamos la $(n-2)$-ésima, multiplicada también por a_1, etc., finalmente, de la segunda fila restamos la primera, multiplicada por a_1. Obtenemos:

$$d = \begin{vmatrix} 1 & 1 & 1 & \dots & 1 \\ 0 & a_2 - a_1 & a_3 - a_1 & \dots & a_n - a_1 \\ 0 & a_2^2 - a_1 a_2 & a_3^2 - a_1 a_3 & \dots & a_n^2 - a_1 a_n \\ \dots & \dots & \dots & \dots & \dots \\ 0 & a_2^{n-1} - a_1 a_2^{n-2} & a_3^{n-1} - a_1 a_3^{n-2} & \dots & a_n^{n-1} - a_1 a_n^{n-2} \end{vmatrix}.$$

Desarrollando este determinante por los elementos de la primera columna, llegamos a un determinante de $(n-1)$-ésimo orden; después de sacar fuera del determinante todos los factores comunes de todas las columnas, éste toma la forma:

$$d = (a_2 - a_1)(a_3 - a_1) \dots (a_n - a_1) \cdot \begin{vmatrix} 1 & 1 & \dots & 1 \\ a_2 & a_3 & \dots & a_n \\ a_2^2 & a_3^2 & \dots & a_n^2 \\ \dots & \dots & \dots & \dots \\ a_2^{n-2} & a_3^{n-2} & \dots & a_n^{n-2} \end{vmatrix}.$$

El último factor es el determinante de Vandermonde de $(n-1)$-ésimo orden que, por la suposición hecha, es igual al producto de todas las diferencias $a_i - a_j$ para $2 \leqslant j < i \leqslant n$. Por consiguiente, empleando el símbolo Π para indicar el producto, se puede escribir:

$$d = (a_2 - a_1)(a_3 - a_1) \dots (a_n - a_1) \prod_{2 \leqslant j < i \leqslant n} (a_i - a_j) = \prod_{1 \leqslant j < i \leqslant n} (a_i - a_j).$$

Del mismo modo se puede demostrar que el determinante

$$d' = \begin{vmatrix} a_1^{n-1} & a_2^{n-1} & a_3^{n-1} & \dots & a_n^{n-1} \\ \dots & \dots & \dots & \dots & \dots \\ a_1^2 & a_2^2 & a_3^2 & \dots & a_n^2 \\ a_1 & a_2 & a_3 & \dots & a_n \\ 1 & 1 & 1 & \dots & 1 \end{vmatrix}$$

es igual al producto de todas las diferencias posibles $a_i - a_j$, donde $1 \leqslant i < j \leqslant n$, es decir,

$$d' = \prod_{1 \leqslant i < j \leqslant n} (a_i - a_j).$$

Generalizando los desarrollos del determinante por los elementos de una fila o columna, obtenidos anteriormente, demostraremos el siguiente teorema del *desarrollo del determinante por los menores de unas cuantas filas o columnas.*

Teorema de Laplace. *Supongamos que en un determinante d de orden n se han elegido arbitrariamente k filas (o k columnas), $1 \leqslant k \leqslant n-1$. Entonces, la suma de los productos de todos los menores de k-ésimo orden, contenidos en las filas elegidas, por sus complementos algebraicos es igual al determinante d.*

Demostración. Supongamos que en el determinante d se han elegido las filas, cuyos números de orden son $i_1, i_2, \ldots, i_k$. Sabemos que el producto de cualquier menor M de k-ésimo orden, situado en estas filas, por su complemento algebraico consta de cierta cantidad de términos del determinante d, tomados con los mismos signos que tenían en el determinante. Por consiguiente, el teorema quedará demostrado, si demostramos que haciendo recorrer a M todos los menores de k-ésimo orden, situados en las filas elegidas, obtenemos todos los términos del determinante, no encontrándose ninguno de ellos dos veces.

Sea

$$a_{1\alpha_1}a_{2\alpha_2}\ldots a_{n\alpha_n} \tag{3}$$

un término arbitrario del determinante d. Tomemos aparte el producto de los elementos de este término, pertenecientes a las filas elegidas, y cuyos números de orden son $i_1, i_2, \ldots, i_k$. Esto será el producto

$$a_{i_1\alpha_{i_1}}a_{i_2\alpha_{i_2}}\ldots a_{i_k\alpha_{i_k}}; \tag{4}$$

k factores de este producto están en k columnas diferentes, precisamente en las columnas con los números de orden $\alpha_{i_1}, \alpha_{i_2}, \ldots, \alpha_{i_k}$. Por consiguiente, estos números de orden de las columnas se determinan por el término (3). Si designamos con M el menor de k-ésimo orden, situado en la intersección de las columnas que tienen estos números de orden $\alpha_{i_1}, \alpha_{i_2}, \ldots, \alpha_{i_k}$, y de las filas elegidas anteriormente, con los números de orden $i_1, i_2, \ldots, i_k$, el producto (4) será uno de *los términos del menor M. El producto de todos los elementos* del término (3), no incluidos en (4), será un término de su menor complementario. Por lo tanto, todo término del determinante forma parte del producto de un menor determinado de k-ésimo orden situado en las filas elegidas por su menor complementario, y además es un producto de unos términos determinados de estos dos menores. Finalmente, para obtener el término tomado del determinante, con el mismo signo que tiene en el determinante, no queda más que susti-

tuir el menor complementario por el complemento algebraico. Con esto termina la demostración del teorema.

Se podía haber demostrado el teorema de otro modo. A saber: el producto de cualquier menor M de k-ésimo orden, situado en las filas elegidas, por su complemento algebraico, consta de $k!$ $(n-k)!$ términos. Esto es debido a que el menor M de k-ésimo orden se compone de $k!$ términos, y su complemento algebraico, diferenciándose posiblemente solamente en el signo del menor de orden $n-k$, contiene $(n-k)!$ términos. Por otra parte, el número de menores de k-ésimo orden, contenidos en las filas que hemos elegido, es igual al número de combinaciones de n sobre k, es decir, es igual al número

$$\frac{n!}{k!\,(n-k)!}.$$

Multiplicando, obtenemos que la suma de los productos de todos los menores de k-ésimo orden de las filas elegidas, por sus complementos algebraicos, consta de $n!$ sumandos. Sin embargo, éste es también, el número total de términos del determinante d. Por consiguiente, el teorema quedará demostrado, si demostramos que cualquier término del determinante d está incluido por lo menos una vez (y entonces, será una vez, exactamente) en la suma considerada de productos de menores por sus complementos algebraicos. Para esto no le queda más al lector que repetir (con ciertas simplificaciones) los razonamientos expuestos en la demostración precedente.

El teorema de Laplace permite reducir el cálculo de un determinante de n-ésimo orden al cálculo de unos cuantos determinantes de órdenes k y $n-k$. Resultará que habrá muchos determinantes nuevos de éstos y, por lo tanto, tiene sentido aplicar el teorema de Laplace solamente en el caso en que se puedan elegir en el determinante k filas (o columnas), de modo que muchos de los menores de k-ésimo orden situados en estas filas sean iguales a cero.

Ejemplos.

1. Sea dado un determinante, cuyos elementos situados en las primeras k filas y últimas $n-k$ columnas son iguales a cero:

$$d=\begin{vmatrix} a_{11} & \dots & a_{1k} & & & \\ \cdot & \cdot & \cdot & & 0 & \\ a_{k1} & \dots & a_{kk} & & & \\ a_{k+1,1} & \dots & a_{k+1,k} & a_{k+1,k+1} & \dots & a_{k+1,n} \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ a_{n1} & \dots & a_{nk} & a_{n,k+1} & \dots & a_{nn} \end{vmatrix}.$$

Este determinante es igual al producto de dos de sus menores:

$$d=\begin{vmatrix} a_{11} & \dots & a_{1k} \\ \cdot & \cdot & \cdot \\ a_{k1} & \dots & a_{kk} \end{vmatrix}\cdot\begin{vmatrix} a_{k+1,k+1} & \dots & a_{k+1,n} \\ \cdot & \cdot & \cdot \\ a_{n,k+1} & \dots & a_{nn} \end{vmatrix}.$$

Para la demostración es suficiente desarrollar el determinante por los menores de las primeras k filas.

2. Sea dado un determinante d de orden $2n$, en cuyo ángulo superior de la izquierda figura un menor formado totalmente por ceros. Si los menores de n-ésimo orden, situados en los ángulos superior de la derecha, inferior de la izquierda e inferior de la derecha del determinante, se designan con M, M' y M'' respectivamente, es decir, que el determinante d se puede escribir simbólicamente en la forma $d = \begin{vmatrix} O & M \\ M' & M'' \end{vmatrix}$, entonces, $d = (-1)^n \, MM'$.

Para la demostración, desarrollamos el determinante por las primeras n filas y observamos que

$$s_M = (1+2+\ldots+n)+[(n+1)+(n+2)\ldots+2n] = n+2n^2,$$

es decir, s_M y n tienen una misma paridad.

3. Calcular el determinante

$$d = \begin{vmatrix} -4 & 1 & 2 & -2 & 1 \\ 0 & 3 & 0 & 1 & -5 \\ 2 & -3 & 1 & -3 & 1 \\ -1 & -1 & 3 & -1 & 0 \\ 0 & 4 & 0 & 2 & 5 \end{vmatrix}.$$

Desarrollándolo por los menores de la primera y tercera columnas, que contienen ceros colocados adecuadamente, obtenemos:

$$d = (-1)^{1+3+1+3}\begin{vmatrix} -4 & 2 \\ 2 & 1 \end{vmatrix}\cdot\begin{vmatrix} 3 & 1 & -5 \\ -1 & -1 & 0 \\ 4 & 2 & 5 \end{vmatrix} +$$

$$+ (-1)^{1+4+1+3}\begin{vmatrix} -4 & 2 \\ -1 & 3 \end{vmatrix}\cdot\begin{vmatrix} 3 & 1 & -5 \\ -3 & -3 & 1 \\ 4 & 2 & 5 \end{vmatrix} +$$

$$+ (-1)^{3+4+1+3}\begin{vmatrix} 2 & 1 \\ -1 & 3 \end{vmatrix}\cdot\begin{vmatrix} 1 & -2 & 1 \\ 3 & 1 & -5 \\ 4 & 2 & 5 \end{vmatrix} =$$

$$= (-8)\cdot(-20)-(-10)\cdot(-62)-7\cdot 87 = -1069.$$

§ 7. Regla de Cramer

La teoría de los determinantes de n-ésimo orden expuesta anteriormente, permite mostrar que estos determinantes, introducidos solamente por analogía con los determinantes de segundo y tercer orden, pueden ser utilizados del mismo modo que estos últimos para la resolución de sistemas de ecuaciones lineales. Sin embargo, primero haremos una observación complementaria, ligada con los desarrollos de los determinantes por los elementos de una fila o columna; en adelante, esta observación va a ser empleada a menudo.

Desarrollemos el determinante

$$d=\begin{vmatrix} a_{11} \dots a_{1j} \dots a_{1n} \\ a_{21} \dots a_{2j} \dots a_{2n} \\ \dots\dots\dots\dots \\ a_{n1} \dots a_{nj} \dots a_{nn} \end{vmatrix}$$

por la j-ésima columna:

$$d=a_{1j}A_{1j}+a_{2j}A_{2j}+\dots+a_{nj}A_{nj},$$

y sustituyamos después en este desarrollo los elementos de **la** j-ésima columna por el sistema de n números arbitrarios b_1, $b_2, \dots, b_n$. La expresión

$$b_1A_{1j}+b_2A_{2j}+\dots+b_nA_{nj},$$

representa el desarrollo por los elementos de la j-ésima columna del determinante

$$d'=\begin{vmatrix} a_{11} \dots b_1 \dots a_{1n} \\ a_{21} \dots b_2 \dots a_{2n} \\ \dots\dots\dots\dots \\ a_{n1} \dots b_n \dots a_{nn} \end{vmatrix},$$

obtenido del determinante d sustituyendo su j-ésima columna por la columna de los números $b_1, b_2, \dots, b_n$. En efecto, la sustitución de la j-ésima columna del determinante d no afecta a los menores de los elementos de esta columna y, por lo tanto, no afecta a sus complementos algebraicos.

Apliquemos esto al caso en que en lugar de los números $b_1, b_2, \dots, b_n$ se toman los elementos de la k-ésima columna del determinante d para $k \neq j$. El determinante que se obtiene después de esta sustitución contendrá dos columnas iguales (la j-ésima y la k-ésima) y, por eso, será igual a cero. Por consiguiente, será igual a cero también el desarrollo de este determinante por los elementos de su j-ésima columna, es decir,

$$a_{1k}A_{1j}+a_{2k}A_{2j}+\dots+a_{nk}A_{nj}=0 \text{ para } j \neq k.$$

Por lo tanto, la *suma de los productos de todos los elementos de una columna del determinante por los complementos algebraicos de los elementos correspondientes de otra columna es igual a cero.* Naturalmente, este resultado es válido también para las *filas* del determinante.

Pasemos a estudiar los sistemas de ecuaciones lineales. Por ahora nos limitaremos al caso de sistemas en los que el **número de**

ecuaciones es igual al número de incógnitas, o sea, a los sistemas de la forma

$$\left.\begin{array}{l} a_{11}x_1+a_{12}x_2+\ldots+a_{1n}x_n=b_1, \\ a_{21}x_1+a_{22}x_2+\ldots+a_{2n}x_n=b_2, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ a_{n1}x_1+a_{n2}x_2+\ldots+a_{nn}x_n=b_n. \end{array}\right\} \qquad (1)$$

Además, supondremos que el determinante d de los coeficientes de las incógnitas del sistema, denominado abreviadamente *determinante del sistema*, es diferente de cero. En estas condiciones demostraremos que el sistema (1) es compatible e incluso determinado.

En el § 2, al resolver un sistema de tres ecuaciones con tres incógnitas, multiplicábamos cada una de las ecuaciones por cierto factor y después sumábamos estas ecuaciones, resultando iguales a cero los coeficientes de dos de las tres incógnitas. Ahora vemos claramente que los factores que empleábamos eran los complementos algebraicos en el determinante del sistema, del elemento que en la ecuación dada es coeficiente de la incógnita buscada. Este mismo método se va a emplear para la resolución del sistema (1).

Supongamos primero que el sistema (1) es compatible y que $\alpha_1, \alpha_2, \ldots, \alpha_n$ es una de sus soluciones. Por consiguiente, se cumplen las igualdades

$$\left.\begin{array}{l} a_{11}\alpha_1+a_{12}\alpha_2+\ldots+a_{1n}\alpha_n=b_1, \\ a_{21}\alpha_1+a_{22}\alpha_2+\ldots+a_{2n}\alpha_n=b_2, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ a_{n1}\alpha_1+a_{n2}\alpha_2+\ldots+a_{nn}\alpha_n=b_n. \end{array}\right\} \qquad (2)$$

Sea j cualquiera de los números 1, 2, . . ., n. Multipliquemos ambos miembros de la primera de las igualdades (2) por A_{1j}, es decir, por el complemento algebraico del elemento a_{1j} en el determinante d del sistema; ambos miembros de la segunda igualdad, por A_{2j}, etc., y finalmente, ambos miembros de la última, por A_{nj}. Sumando después por separado los primeros miembros y los segunpos miembros de todas estas igualdades, llegamos a la siguiente igualdad:

$$\begin{array}{l} (a_{11}A_{1j}+a_{21}A_{2j}+\ldots+a_{n1}A_{nj})\,\alpha_2+ \\ \qquad +(a_{12}A_{1j}+a_{22}A_{2j}+\ldots+a_{n2}A_{nj})\,\alpha_1+ \\ \qquad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ \qquad +(a_{1j}A_{1j}+a_{2j}A_{2j}+\ldots+a_{nj}A_{nj})\,\alpha_j+ \\ \qquad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ \qquad +(a_{1n}A_{1j}+a_{2n}A_{2j}+\ldots+a_{nn}A_{nj})\,\alpha_n= \\ \qquad\qquad =b_1A_{1j}+b_1A_{2j}+\ldots b_nA_{nj}. \end{array}$$

El coeficiente de α_j en esta igualdad es igual a d, mientras que, en virtud de la observación hecha anteriormente, los coeficientes de los demás α son iguales a cero; el miembro independiente es igual al determinante que se obtiene del determinante d después de sustituir en él la j-ésima columna por la columna de los términos independientes del sistema (1). Si designamos este último determinante, igual que en el § 2, con d_j, nuestra igualdad toma la forma

$$d\alpha_j = d_j,$$

de donde

$$\alpha_j = \frac{d_j}{d},$$

puesto que $d \neq 0$. De este modo, queda demostrado que si el sistema (1) es compatible, éste posee solución única:

$$\alpha_1 = \frac{d_1}{d}, \ \alpha_2 = \frac{d_2}{d}, \ \ldots, \ \alpha_n = \frac{d_n}{d}. \tag{3}$$

Demostremos ahora que el sistema de números (3) satisface realmente al sistema de ecuaciones (1), es decir, que el sistema (1) es compatible. A continuación emplearemos las siguientes notaciones muy usuales.

Toda suma de la forma $a_1 + a_2 + \ldots + a_n$ se indicará abreviadamente mediante $\sum_{i=1}^{n} a_i$. Si se considera una suma, cuyos sumandos a_{ij} están provistos de dos subíndices, siendo $i = 1, 2, \ldots, n$, $j = 1, 2, \ldots, m$, se pueden tomar primero las sumas de elementos con el primer subíndice fijado, o sea, las sumas $\sum_{j=1}^{m} a_{ij}$, donde $i = 1, 2, \ldots, n$, y después, sumar todas estas sumas. Entonces, para la suma de todos los elementos a_{ij}, obtenemos la expresión

$$\sum_{i=1}^{n} \sum_{j=1}^{m} a_{ij}.$$

No obstante, se podrían sumar primero los sumandos a_{ij} con el segundo subíndice fijado y sumar después las sumas obtenidas. Por lo tanto,

$$\sum_{i=1}^{n} \sum_{j=1}^{m} a_{ij} = \sum_{j=1}^{m} \sum_{i=1}^{n} a_{ij},$$

o sea, en la *suma doble se puede cambiar el orden de los sumandos.*

Pongamos ahora en la i-ésima ecuación del sistema (1) los valores (3) de las incógnitas. Como el primer miembro de la i-ésima

ecuación se puede escribir de la forma $\sum_{j=1}^{n} a_{ij}x_j$ y como $d_j = \sum_{k=1}^{n} b_k A_{kj}$, obtenemos:

$$\sum_{j=1}^{n} a_{ij} \cdot \frac{d_j}{d} = \frac{1}{d} \sum_{j=1}^{n} a_{ij} \left(\sum_{k=1}^{n} b_k A_{kj} \right) = \frac{1}{d} \sum_{k=1}^{n} b_k \left(\sum_{j=1}^{n} a_{ij} A_{kj} \right).$$

Respecto a estas transformaciones, observemos que el número $\frac{1}{d}$ es un factor común de todos los sumandos, por lo cual, se le ha sacado fuera de la suma; además, después de haber cambiado el orden de los sumandos, el factor b_k se ha sacado fuera de la suma interior, ya que no depende del subíndice j de la suma interior.

Ya sabemos que la expresión $\sum_{j=1}^{n} a_{ij}A_{kj} = a_{i1}A_{k1} + a_{i2}A_{k2} + \ldots + a_{in}A_{kn}$ es igual a d para $k = i$, e igual a 0 para los demás k. Por lo tanto, en nuestra suma exterior respecto a k quedará un sumando, precisamente $b_i d$:

$$\sum_{j=1}^{n} a_{ij} \cdot \frac{d_j}{d} = \frac{1}{d} \cdot b_i d = b_i.$$

De este modo, queda demostrado que el sistema de números (3) es, verdaderamente, solución del sistema de ecuaciones (1).

Hemos obtenido el siguiente resultado importante:

Un sistema de n ecuaciones lineales con n incógnitas, cuyo determinante es diferente de cero, tiene solución, la cual, además, es única. Esta solución se obtiene por las fórmulas (3), es decir, por la *regla de Cramer*; la formulación de esta regla es igual que en el caso de un sistema de dos ecuaciones (véase § 2).

Ejemplo. *Resolver el sistema de ecuaciones lineales*

$$\left.\begin{aligned} 2x_1 + x_2 - 5x_3 + x_4 &= 8, \\ x_1 - 3x_2 \qquad - 6x_4 &= 9, \\ 2x_2 - x_3 + 2x_4 &= -5, \\ x_1 + 4x_2 - 7x_3 + 6x_4 &= 0. \end{aligned}\right\}$$

El determinante de este sistema es diferente de cero:

$$d = \begin{vmatrix} 2 & 1 & -5 & 1 \\ 1 & -3 & 0 & -6 \\ 0 & 2 & -1 & 2 \\ 1 & 4 & -7 & 6 \end{vmatrix} = 27,$$

por lo que se puede aplicar al sistema la regla de Cramer. Los valores de las incógnitas tendrán en los numeradores los determinantes

$$d_1=\begin{vmatrix} 8 & 1 & -5 & 1 \\ 9 & -3 & 0 & -6 \\ -5 & 2 & -1 & 2 \\ 0 & 4 & -7 & 6 \end{vmatrix}=81, \quad d_2=\begin{vmatrix} 2 & 8 & -5 & 1 \\ 1 & 9 & 0 & -6 \\ 0 & -5 & -1 & 2 \\ 1 & 0 & -7 & 6 \end{vmatrix}=-108,$$

$$d_3=\begin{vmatrix} 2 & 1 & 8 & 1 \\ 1 & -3 & 9 & -6 \\ 0 & 2 & -5 & 2 \\ 1 & 4 & 0 & 6 \end{vmatrix}=-27, \quad d_4=\begin{vmatrix} 2 & 1 & -5 & 8 \\ 1 & -3 & 0 & 9 \\ 0 & 2 & -1 & -5 \\ 1 & 4 & -7 & 0 \end{vmatrix}=27.$$

Por lo tanto

$$x_1=3, \ x_2=-4, \ x_3=-1, \ x_4=1$$

será la solución de nuestro sistema *y*, además, la única.

Hemos excluido el caso en que el determinante del sistema de n ecuaciones lineales con n incógnitas (1) es igual a cero. Este caso lo dejamos para el cap. 2, donde hallará su sitio en la teoría general de los sistemas de cualquier número de ecuaciones con cualquier número de incógnitas.

Referente a los sistemas de n ecuaciones lineales con n incógnitas, haremos otra observación más. Sea dado un sistema de n ecuaciones lineales *homogéneas* con n incógnitas (véase el § 1):

$$\left.\begin{array}{l} a_{11}x_1+a_{12}x_2+\ldots+a_{1n}x_n=0, \\ a_{21}x_1+a_{22}x_2+\ldots+a_{2n}x_n=0, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ a_{n1}x_1+a_{n2}x_2+\ldots+a_{nn}x_n=0. \end{array}\right\} \tag{4}$$

En este caso, todos los determinates d_j, $j=1, 2, \ldots, n$, contienen una columna formada por ceros y, por eso, son iguales a cero. Por lo tanto, si el determinante del sistema (4) es diferente de cero, es decir, si a este sistema se le puede aplicar la regla de Cramer, su única solución será la solución nula

$$x_1=0, \ x_2=0, \ \ldots, \ x_n=0. \tag{5}$$

De aquí se desprende la siguiente conclusión:

Si un sistema de n ecuaciones lineales homogéneas con n incógnitas tiene soluciones diferentes de la nula, entonces el determinante de este sistema es necesariamente igual a cero.

En el § 12 se mostrará que, viceversa, si el determinante de un sistema de éstos es igual a cero, además de la solución nula, cuya existencia es evidente para cualquier sistema de ecuaciones homogéneas, tendrán que existir también otras soluciones.

Ejemplo. ¿Para qué valores de k, el sistema de ecuaciones

$$\left.\begin{aligned} kx_1+x_2=0, \\ x_1+kx_2=0 \end{aligned}\right\}$$

puede tener soluciones no nulas?

El determinante de este sistema

$$\begin{vmatrix} k & 1 \\ 1 & k \end{vmatrix} = k^2-1$$

será igual a cero solamente para $k = \pm 1$. *Es fácil comprobar que para cada* uno de estos dos valores de k, el sistema dado posee verdaderamente soluciones diferentes de la nula.

La importancia de la regla de Cramer consiste fundamentalmente *en que, en los casos en que es aplicable esta regla, ésta da una expresión* explícita para la solución del sistema mediante los coeficientes del mismo. Sin embargo, la aplicación práctica de la regla de Cramer va aparejada con cálculos muy complicados: en el caso de un sistema *de n ecuaciones lineales con n incógnitas, se tienen que calcular* $n+1$ determinantes de n-ésimo orden. El método de eliminación sucesiva de las incógnitas, expuesto en el § 1, es en este sentido mucho más cómodo, puesto que los cálculos que se necesitan para aplicar este método son, en esencia, equivalentes a los que se tienen que realizar al calcular *un solo* determinante de n-ésimo orden.

En algunas aplicaciones aparecen sistemas de ecuaciones lineales cuyos coeficientes y términos independientes son números reales, obtenidos al hacer mediciones de algunas cantidades físicas, es decir, que se conocen sólo aproximadamente, con cierta exactitud. A veces, los métodos expuestos anteriormente para la resolución de tales sistemas son inadecuados, debido a que proporcionan resultados poco exactos. En su lugar, se han elaborado diversos *métodos de iteración*, o sea, métodos que permiten resolver los sistemas indicados de ecuaciones mediante una aproximación sucesiva de las incógnitas. La exposición de estos métodos puede consultarla el lector en las obras sobre la teoría de las aproximaciones.

CAPITULO II

SISTEMAS DE ECUACIONES LINEALES (TEORIA GENERAL)

§ 8. Espacio vectorial de n dimensiones

Para la elaboración de la teoría general de los sistemas de ecuaciones lineales no es suficiente el aparato construido que nos sirvió satisfactoriamente para la resolución de los sistemas en que se puede aplicar la regla de Cramer. Además de los determinantes y las matrices tenemos que utilizar un nuevo concepto que, posiblemente, sea de mayor interés para la matemática en general: el concepto de *espacio vectorial de varias dimensiones.*

Hagamos primero unas cuantas observaciones previas. Por el curso de geometría analítica se sabe que todo punto en el plano se determina (dados los ejes coordenados) por sus dos coordenadas, o sea, por un sistema ordenado de dos números reales; todo vector en el plano se determina por sus dos componentes, o sea, nuevamente, por un sistema ordenado de dos números reales. De modo análogo, todo punto en el espacio de tres dimensiones se determina por sus tres coordenadas, y todo vector en el espacio se determina por sus tres componentes.

En la geometría, y también en la mecánica y en la física, se suelen estudiar frecuentemente algunos objetos, para cuya determinación no son suficientes tres números reales. Veamos, por ejemplo, el conjunto de las esferas en el espacio. Para que la esfera esté determinada por completo, es necesario que estén dadas las coordenadas de su centro y el radio, o sea, hay que señalar un sistema ordenado de cuatro números reales, de los cuales el último (el radio) sólo puede tomar, a su vez, valores positivos. Examinemos, por otra parte, las diferentes posiciones de un cuerpo sólido en el espacio. La posición del cuerpo quedará determinada por completo, si se indican las coordenadas de su centro de gravedad (o sea, tres números reales), la dirección de un eje fijo que pase por el centro de gravedad (dos números: dos, de los tres cosenos directores) y, por fin, el ángulo de rotación alrededor de este eje. Por lo tanto, la posición de un sólido en el espacio se determina por un sistema ordenado de seis números reales.

Estos ejemplos nos sugieren la oportunidad de estudiar el conjunto de todos los sistemas ordenados posibles de n números reales. Precisamente este conjunto, después de haber introducido en él las operaciones de adición y multiplicación (cosa que se hará a continuación por analogía con las operaciones correspondientes sobre los vectores del espacio tridimensional, expresadas mediante las componentes), se denomina espacio vectorial de n dimensiones. Por consiguiente, el espacio de n dimensiones es solamente una formación algebraica que conserva ciertas propiedades elementales del conjunto de los vectores del espacio de tres dimensiones, que parten del origen de coordenadas.

Un sistema ordenado de n números

$$\alpha = (a_1, a_2, \ldots, a_n) \tag{1}$$

se llama *vector de n dimensiones*. Los números a_i, $i = 1, 2, \ldots, n$, se denominarán *componentes* del vector α. Se dirá que los vectores α y

$$\beta = (b_1, b_2, \ldots, b_n) \tag{2}$$

son *iguales*, si coinciden sus componentes situadas en lugares iguales, o sea, si $a_i = b_i$ para $i = 1, 2, \ldots, n$. Para designar los vectores se emplearán en adelante las letras griegas minúsculas, mientras que las letras latinas minúsculas se utilizarán para designar los números.

Como ejemplos de vectores, señalemos los siguientes: 1) Los vectores-segmentos que parten del origen de coordenadas, en el plano o en el espacio de tres dimensiones, estando fijado el sistema de coordenadas, serán vectores de dos y tres dimensiones, respectivamente, en el sentido de la definición dada anteriormente. 2) Los coeficientes de cualquier ecuación lineal con n incógnitas forman un vector de n dimensiones. 3) Toda solución de cualquier sistema de ecuaciones lineales con n incógnitas es un vector de n dimensiones. 4) Dada una matriz de s filas y n columnas, sus filas son vectores de n dimensiones y sus columnas, vectores de s dimensiones. 5) La misma matriz de s filas y n columnas se puede considerar como un vector de sn dimensiones: es suficiente leer seguidamente los elementos de la matriz, fila por fila; en particular, toda matriz cuadrada de orden n se puede considerar como un vector de n^2 dimensiones. Es evidente, además, que cualquier vector de n^2 dimensiones se puede obtener de este modo de una matriz cuadrada de orden n.

Se llama *suma* de los vectores (1) y (2) al vector

$$\alpha + \beta = (a_1 + b_1,\ a_2 + b_2,\ \ldots,\ a_n + b_n), \tag{3}$$

cuyas componentes son iguales a las sumas de las componentes correspondientes de los vectores que se suman. La adición de vectores

está sujeta a las leyes conmutativa y asociativa, puesto que la adición de los números está sujeta a estas leyes.

El *vector nulo* desempeña el papel de cero

$$0=(0,\ 0,\ \ldots,\ 0).$$

En efecto

$$\alpha+0=(a_1+0, a_2+0,\ \ldots,\ a_n+0)=(a_1,\ a_2,\ \ldots,\ a_n)=\alpha.$$

Para designar el vector nulo emplearemos el mismo símbolo 0 que se emplea para el número cero; nunca encontraremos dificultad alguna para averiguar si en el momento dado se trata del número cero o del vector nulo; sin embargo, al estudiar los próximos párrafos, el lector tiene que recordar que el símbolo 0 se puede emplear en diversos sentidos.

El vector

$$-\alpha=(-a_1,\ -a_2,\ \ldots,\ -a_n). \tag{5}$$

se denominará vector *opuesto* del vector (1). Es evidente, que $\alpha+(-\alpha)=0$. Ahora, es fácil demostrar que para la adición de vectores existe la operación inversa: la sustracción; la *diferencia* de los vectores (1) y (2) es el vector $\alpha-\beta=\alpha+(-\beta)$, o sea,

$$\alpha-\beta=(a_1-b_1,\ a_2-b_2,\ \ldots,\ a_n-b_n). \tag{6}$$

La suma de vectores de n dimensiones, definida por la fórmula (3), fue originada por la suma geométrica de vectores en el plano o en el espacio de tres dimensiones, efectuada de acuerdo a la regla del paralelogramo. En la geometría se define también el producto de un vector por un número real (por un «escalar»): multiplicar el vector α por el número k significa, siendo $k>0$, que el vector α se alarga k veces (o que se contrae, si $k<1$), y siendo $k<0$, que se alarga $|k|$ veces y se cambia su dirección por la opuesta. Expresando esta regla mediante las componentes del vector y pasando al caso general considerado, obtenemos la definición siguiente:

Se llama *producto del vector* (1) *por el número* k, al vector

$$k\alpha=\alpha k=(ka_1,\ ka_2,\ \ldots,\ ka_n), \tag{7}$$

cuyas componentes son iguales al producto de las correspondientes *componentes del vector* α *por el número* k.

De esta definición se deducen las siguientes importantes propiedades, cuyas demostraciones se dejan al lector:

$$k(\alpha\pm\beta)=k\alpha\pm k\beta; \tag{8}$$

$$(k\pm l)\alpha=k\alpha\pm l\alpha; \tag{9}$$

$$k(l\alpha)=(kl)\alpha; \tag{10}$$

$$1\cdot\alpha=\alpha. \tag{11}$$

Con la misma facilidad se comprueban, aunque pueden obtenerse también como consecuencia de las propiedades (8) — (11), las propiedades siguientes:

$$0 \cdot \alpha = 0; \qquad (12)$$

$$(-1) \cdot \alpha = -\alpha; \qquad (13)$$

$$k \cdot \mathbf{0} = \mathbf{0}; \qquad (14)$$

$$\text{si } k\alpha = \mathbf{0}, \text{ entonces } k = 0, \text{ o bien } \alpha = \mathbf{0}. \qquad (15)$$

El conjunto de todos los vectores de n dimensiones con componentes reales, considerado junto con las operaciones de suma de vectores y de multiplicación de un vector por un número, determinadas en el mismo, se llama *espacio vectorial de n dimensiones*.

Subrayemos que en la definición de espacio vectorial de n dimensiones no está incluida ninguna multiplicación de un vector por otro vector. Sería fácil definir el producto de vectores: se podría suponer, por ejemplo, que las componentes del producto de vectores fuesen iguales a los productos de las componentes correspondientes de los factores. Sin embargo, una tal multiplicación no tendría aplicaciones serias. Así, pues, los segmentos-vectores que parten del origen de coordenadas, en el plano o en el espacio de tres dimensiones, (se supone que se ha fijado un sistema de coordenadas), forman un espacio vectorial de dos y de tres dimensiones, respectivamente. Como se ha señalado anteriormente, en este ejemplo, la suma de vectores y el producto de un vector por un número tienen un sentido geométrico importante, mientras que al producto de vectores definido mediante la multiplicación de sus componentes no se le puede dar ninguna significación geométrica racional.

Veamos otro ejemplo más. El primer miembro de una ecuación lineal con n incógnitas, es decir, la expresión de la forma

$$f = a_1x_1 + a_2x_2 + \ldots + a_nx_n,$$

se llama *forma lineal* en las incógnitas $x_1, x_2, \ldots, x_n$. Es evidente que la forma lineal f queda completamente determinada por el vector $(a_1, a_2, \ldots, a_n)$ de sus coeficientes; recíprocamente, todo vector n-dimensional determina unívocamente una forma lineal. La suma de vectores y el producto de un vector por un número se convierten en las operaciones correspondientes con las formas lineales; estas operaciones fueron empleadas eficazmente por nosotros en el § 1. La multiplicación de los vectores definida mediante el producto de sus componentes, no tiene tampoco en este ejemplo ningún sentido.

§ 9. Dependencia lineal de vectores

Se dice que el vector β, de un espacio vectorial de n dimensiones, es *proporcional* al vector α, si existe un número k tal que $\beta = k\alpha$ (véase la fórmula (7) del párrafo anterior). En particular, el vector nulo es proporcional a cualquier vector α, debido a la igualdad $0 = 0\cdot\alpha$. Si $\beta = k\alpha$ y $\beta \neq 0$, de donde $k \neq 0$, entonces $\alpha = k^{-1}\beta$; es decir, para los vectores no nulos, la proporcionalidad posee la propiedad de simetría.

Una generalización del concepto de proporcionalidad de vectores es la noción siguiente (con la que ya nos encontramos en el § 4, para el caso de las filas de las matrices): se dice que el vector β es una *combinación lineal* de los vectores $\alpha_1, \alpha_2, \ldots, \alpha_s$, si existen unos números $l_1, l_2, \ldots, l_s$ tales que

$$\beta = l_1\alpha_1 + l_2\alpha_2 + \ldots + l_s\alpha_s.$$

Por lo tanto, la j-ésima componente del vector β, $j = 1, 2, \ldots, n$, en virtud de la definición de la suma de vectores y del producto de un vector por un número, es igual a la suma de los productos de las j-ésimas componentes de los vectores $\alpha_1, \alpha_2, \ldots, \alpha_s$ por los números $l_1, l_2, \ldots, l_s$ correspondientemente.

Se dice que el sistema de vectores

$$\alpha_1, \alpha_2, \ldots, \alpha_{r-1}, \alpha_r \quad (r \geqslant 2) \tag{1}$$

es *linealmente dependiente*, si al menos uno de estos vectores puede expresarse como combinación lineal de los demás vectores del sistema (1); en caso contrario, se dice que el sistema (1) es *linealmente independiente*.

Señalemos otra forma de esta importantísima definición: el sistema de vectores (1) es linealmente dependiente, si existen unos números $k_1, k_2, \ldots, k_r$, *entre los cuales al menos uno es diferente de cero*, de modo que se verifica la igualdad

$$k_1\alpha_1 + k_2\alpha_2 + \ldots + k_r\alpha_r = 0. \tag{2}$$

La demostración de la equivalencia de estas dos definiciones no representa dificultad alguna. Sea, por ejemplo, el vector α_r del sistema (1), combinación lineal de los demás vectores:

$$\alpha_r = l_1\alpha_1 + l_2\alpha_2 + \ldots + l_{r-1}\alpha_{r-1}.$$

De aquí se deduce la igualdad

$$l_1\alpha_1 + l_2\alpha_2 + \ldots + l_{r-1}\alpha_{r-1} - \alpha_r = 0,$$

es decir, una igualdad de la forma (2), donde $k_i = l_i$ para $i = 1, 2, \ldots, r-1$ y $k_r = -1$, es decir, $k_r \neq 0$. Recíprocamente,

supongamos que los vectores (1) están ligados por la relación (2), en la que, por ejemplo, $k_r \neq 0$. Entonces,

$$\alpha_r = \left(-\frac{k_1}{k_r}\right)\alpha_1 + \left(-\frac{k_2}{k_r}\right)\alpha_2 + \ldots + \left(-\frac{k_{r-1}}{k_r}\right)\alpha_{r-1},$$

es decir, resulta que el vector α_r es combinación lineal de los vectores $\alpha_1, \alpha_2, \ldots, \alpha_{r-1}$.

Ejemplo. El sistema de vectores

$$\alpha_1 = (5, 2, 1), \quad \alpha_2 = (-1, 3, 3), \quad \alpha_3 = (9, 7, 5), \quad \alpha_4 = (3, 8, 7)$$

es linealmente dependiente, puesto que los vectores están ligados por la *relación*

$$4\alpha_1 - \alpha_2 - 3\alpha_3 + 2\alpha_4 = 0.$$

En esta relación todos los coeficientes son diferentes de cero. Por otra parte, entre nuestros vectores existen también otras dependencias lineales, en las que algunos de los coeficientes son iguales a cero, por ejemplo

$$2\alpha_1 + \alpha_2 - \alpha_3 = 0, \quad 3\alpha_2 + \alpha_3 - 2\alpha_4 = 0.$$

La segunda de las definiciones de dependencia lineal dada anteriormente, se puede aplicar cuando $r = 1$, o sea, al caso de un sistema compuesto de un solo vector α: *este sistema será linealmente dependiente cuando, y sólo cuando,* $\alpha = 0$. En efecto, si $\alpha = 0$, entonces, por ejemplo, para $k = 1$, se tiene $k\alpha = 0$. Recíprocamente, si $k\alpha = 0$ y $k \neq 0$, entonces, $\alpha = 0$.

Señalemos la siguiente propiedad del concepto de dependencia lineal.

Si un subsistema del sistema de vectores (1) es linealmente dependiente, lo es también todo el sistema (1).

En efecto, supongamos que los vectores $\alpha_1, \alpha_2, \ldots, \alpha_s$ del sistema (1), donde $s < r$, están ligados por la relación

$$k_1\alpha_1 + k_2\alpha_2 + \ldots + k_s\alpha_s = 0,$$

en la que no todos los coeficientes son iguales a cero. De aquí se deduce la relación

$$k_1\alpha_1 + k_2\alpha_2 + \ldots + k_s\alpha_s + 0\cdot\alpha_{s+1} + \ldots + 0\cdot\alpha_r = 0,$$

es decir, el sistema (1) es linealmente dependiente.

De esta propiedad se deduce la *dependencia lineal de cualquier sistema de vectores que contenga dos vectores iguales o, en general, dos vectores proporcionales, así como de cualquier sistema que contenga al vector nulo.* Obsérvese que la propiedad que acabamos de demostrar se puede formular de otra manera: *si el sistema de vectores (1) es linealmente independiente, cualquier subsistema del mismo es también linealmente independiente.*

Aquí surgen las preguntas: ¿puede contener muchos vectores un sistema linealmente independiente de vectores de n dimensiones? y en particular, ¿existen tales sistemas con un número arbitrariamente grande de vectores? Para responder a estas preguntas, consideremos en el espacio vectorial de n dimensiones los vectores

$$\left.\begin{array}{l} \varepsilon_1=(1,\ 0,\ 0,\ \ldots,\ 0), \\ \varepsilon_2=(0,\ 1,\ 0,\ \ldots,\ 0), \\ \cdots\cdots\cdots\cdots \\ \varepsilon_n=(0,\ 0,\ 0,\ \ldots,\ 1), \end{array}\right\} \quad (3)$$

denominados *vectores unitarios* de este espacio.

El sistema de vectores unitarios es linealmente independiente. Sea

$$k_1\varepsilon_1+k_2\varepsilon_2+\ldots+k_n\varepsilon_n=0;$$

como el primer miembro de esta igualdad es igual al vector $(k_1,\ k_2,\ \ldots,\ k_n)$, se tiene

$$(k_1,\ k_2,\ \ldots,\ k_n)=0,$$

o sea, $k_i=0,\ \ i=1,\ 2,\ \ldots,\ n$, puesto que todas las componentes del vector nulo son iguales a cero y la igualdad de vectores es equivalente a la igualdad de sus componentes correspondientes.

Por lo tanto, en el espacio vectorial de n dimensiones hemos hallado un sistema linealmente independiente, compuesto de n vectores. El lector verá más adelante que en realidad, en este espacio existen infinitos sistemas de éstos. Demostremos, por otra parte, el siguiente teorema:

Cualesquiera s vectores del espacio vectorial de n dimensiones forman, para $s>n$, un sistema linealmente dependiente.

En efecto, supongamos que se han dado los vectores

$$\begin{array}{l} \alpha_1=(a_{11},\ a_{12},\ \ldots,\ a_{1n}), \\ \alpha_2=(a_{21},\ a_{22},\ \ldots,\ a_{2n}), \\ \cdots\cdots\cdots\cdots \\ \alpha_s=(a_{s1},\ a_{s2},\ \ldots,\ a_{sn}). \end{array}$$

Tenemos que elegir unos números $k_1,\ k_2,\ \ldots,\ k_s$, no todos iguales a cero, de modo que

$$k_1\alpha_1+k_2\alpha_2+\ldots+k_s\alpha_s=0. \quad (4)$$

Pasando de la igualdad (4) a las igualdades correspondientes entre las componentes, obtenemos

$$\left.\begin{array}{l} a_{11}k_1+a_{21}k_2+\ldots+a_{s1}k_s=0, \\ a_{12}k_1+a_{22}k_2+\ldots+a_{s2}k_s=0, \\ \cdots\cdots\cdots\cdots\cdots \\ a_{1n}k_1+a_{2n}k_2+\ldots+a_{sn}k_s=0. \end{array}\right\} \quad (5)$$

Las igualdades (5) forman, sin embargo, un sistema de n ecuaciones lineales homogéneas respecto a s incógnitas $k_1, k_2, \ldots, k_s$. El número de ecuaciones en este sistema es menor que el número de incógnitas y, por consiguiente, como se ha demostrado al final del § 1, este sistema tiene soluciones no nulas. Por lo tanto, se pueden elegir unos números $k_1, k_2, \ldots, k_s$, no todos iguales a cero, que satisfaga la condición (4). El teorema queda demostrado.

Un sistema linealmente independiente de vectores de n dimensiones

$$\alpha_1, \alpha_2 \ldots, \alpha_r \tag{6}$$

se llamará sistema linealmente independiente, *maximal*, si al agregarle cualquier vector β de n dimensiones, resulta un sistema linealmente dependiente. Como en cualquier dependencia lineal que liga los vectores $\alpha_1, \alpha_2, \ldots, \alpha_r, \beta$, el coeficiente de β tiene que ser diferente de cero (puesto que, en caso contrario, el sistema (6) sería linealmente dependiente), el vector β se expresará linealmente mediante los vectores (6). Por ello, el sistema de vectores (6) es un sistema linealmente independiente *maximal*, cuando, y sólo cuando, los vectores (6) son linealmente independientes, y cualquier vector β de n dimensiones se expresa como combinación lineal de ellos.

De los resultados que hemos obtenido anteriormente se deduce que en el *espacio de n dimensiones, todo sistema, linealmente independiente, compuesto de n vectores, siempre es maximal*, y también, que *cualquier sistema de vectores linealmente independiente maximal no consta de más de n vectores.*

Todo sistema de vectores de n dimensiones, linealmente independiente, está contenido, al menos, en un sistema linealmente independiente maximal. En efecto, si el sistema dado de vectores no es maximal, se le puede agregar un vector de tal modo que el sistema obtenido se mantenga linealmente independiente. Si este sistema nuevo no es todavía maximal, se le puede agregar otro vector más, etc. Naturalmente, este proceso no se puede continuar indefinidamente, puesto que cualquier sistema de vectores de n dimensiones, compuesto de $n + 1$ vectores, es ya linealmente dependiente.

Como cualquier sistema que consta de un sólo vector no nulo es linealmente independiente, resulta que *cualquier vector no nulo está contenido en un sistema linealmente independiente maximal.* Por consiguiente, en el *espacio vectorial de n dimensiones existe una infinidad de diversos sistemas de vectores linealmente independientes maximales.*

Surge la pregunta: ¿existen en este espacio sistemas linealmente independientes maximales que contengan menos de n vectores, o el número de vectores en cualquier sistema de éstos tiene que ser, indispensablemente, igual a n? La respuesta a esta importante pre-

gunta se dará un poco más adelante, después de hacer algunas observaciones.

Se dice frecuentemente, que el vector β se expresa *linealmente mediante el sistema de vectores*

$$\alpha_1, \alpha_2, \ldots, \alpha_r, \tag{7}$$

si β es una combinación lineal de ellos. Se comprende que, si el vector β se expresa linealmente mediante un subsistema de este sistema, entonces se expresa también linealmente mediante el sistema (7). Para demostrar esto, es suficiente tomar los otros vectores con los coeficientes iguales a cero. Generalizando esta terminología, se dice que el *sistema de vectores*

$$\beta_1, \beta_2, \ldots, \beta_s \tag{8}$$

se expresa *linealmente mediante el sistema* (7), si cada vector β_i, $i = 1, 2, \ldots, s$ es combinación lineal de los vectores del sistema (7).

Demostremos que para este concepto se cumple la ley transitiva: *si el sistema* (8) *se expresa linealmente mediante el sistema* (7), *y el sistema de vectores*

$$\gamma_1, \gamma_2, \ldots, \gamma_t \tag{9}$$

se expresa linealmente mediante el sistema (8), *entonces el sistema* (9) *también se expresa linealmente mediante el sistema* (7).

En efecto

$$\gamma_j = \sum_{i=1}^{s} l_{ji}\beta_i, \quad j = 1, 2, \ldots, t, \tag{10}$$

pero $\beta_i = \sum_{m=1}^{r} k_{im}\alpha_m$, $i = 1, 2, \ldots, s$. Sustituyendo en (10) estas expresiones, obtenemos:

$$\gamma_j = \sum_{i=1}^{s} l_{ji}\left(\sum_{m=1}^{r} k_{im}\alpha_m\right) = \sum_{m=1}^{r}\left(\sum_{i=1}^{s} l_{ji}k_{im}\right)\alpha_m,$$

o sea, cualquier vector γ_j, $j = 1, 2, \ldots, t$ es combinación lineal de los vectores del sistema (7).

Dos sistemas de vectores se llaman *equivalentes*, si cada uno de ellos se expresa linealmente mediante el otro. De la ley transitiva que acabamos de demostrar, a la que satisface la propiedad de los sistemas de vectores de expresarse linealmente entre sí, se deduce el cumplimiento de la misma ley para el concepto de equivalencia de los sistemas de vectores. De aquí también se deduce la afirmación siguiente: *siendo equivalentes dos sistemas de vectores, si un vector se expresa linealmente mediante uno de estos sistemas, entonces se expresa también linealmente mediante el otro.*

No se puede afirmar que siendo linealmente independiente uno de dos sistemas de vectores, equivalentes entre sí, lo es también

el otro. Si ambos sistemas son linealmente independientes, se puede enunciar una proposición importante sobre el número de vectores que forman parte de ellos. Pero, demostremos primero el siguiente teorema, que debido al papel que va a desempeñar en adelante, lo denominaremos **teorema fundamental.**

Si en el espacio vectorial de n dimensiones se han dado dos sistemas de vectores:

(I) $$\alpha_1,\ \alpha_2,\ \ldots,\ \alpha_r,$$

(II) $$\beta_1,\ \beta_2,\ \ldots,\ \beta_s,$$

el primero de los cuales es linealmente independiente y se expresa linealmente mediante el segundo, entonces el número de vectores del primer sistema no es superior al número de vectores del segundo sistema, es decir, $r \leqslant s$.

En efecto, supongamos que $r > s$. Por la hipótesis, cada vector del sistema (I) se expresa linealmente mediante el sistema (II):

$$\left.\begin{array}{l} \alpha_1 = \alpha_{11}\beta_1 + \alpha_{12}\beta_2 + \ldots + \alpha_{1s}\beta_s, \\ \alpha_2 = \alpha_{21}\beta_1 + \alpha_{22}\beta_2 + \ldots + \alpha_{2s}\beta_s, \\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot \\ \alpha_r = \alpha_{r1}\beta_1 + \alpha_{r2}\beta_2 + \ldots + \alpha_{rs}\beta_s. \end{array}\right\} \quad (11)$$

Los coeficientes de estas expresiones lineales forman un sistema de r vectores de s dimensiones:

$$\begin{array}{l} \gamma_1 = (a_{11},\ a_{12},\ \ldots,\ a_{1s}), \\ \gamma_2 = (a_{21},\ a_{22},\ \ldots,\ a_{2s}), \\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot \\ \gamma_r = (a_{r1},\ a_{r2},\ \ldots,\ a_{rs}). \end{array}$$

Como $r > s$, estos vectores son linealmente dependientes, o sea,

$$k_1\gamma_1 + k_2\gamma_2 + \ldots + k_r\gamma_r = 0,$$

donde no todos los coeficientes $k_1,\ k_2,\ \ldots,\ k_r$ son iguales a cero. De aquí, llegamos a las siguientes igualdades entre las componentes:

$$\sum_{i=1}^{r} k_i a_{ij} = 0,\ j = 1,\ 2,\ \ldots,\ s. \quad (12)$$

Consideremos ahora la siguiente combinación lineal de los vectores del sistema (I):

$$k_1\alpha_1 + k_2\alpha_2 + \ldots + k_r\alpha_r$$

o, abreviadamente, $\sum_{i=1}^{r} k_i\alpha_i$. Aplicando (11) y (12), resulta:

$$\sum_{i=1}^{r} k_i\alpha_i = \sum_{i=1}^{r} k_i\left(\sum_{j=1}^{s} a_{ij}\beta_j\right) = \sum_{j=1}^{s}\left(\sum_{i=1}^{r} k_i a_{ij}\right)\beta_j = 0;$$

lo que, sin embargo, contradice a la independencia lineal del sistema (I).

Del teorema fundamental que acabamos de demostrar se deduce el resultado siguiente:

Dos sistemas equivalentes de vectores cualesquiera, linealmente independientes, contiene el mismo número de vectores.

Es evidente que dos sistemas maximales cualesquiera de vectores de n dimensiones linealmente independientes, son equivalentes. Por consiguiente, se componen de un mismo número de vectores, y como existen sistemas de este género compuestos de n vectores, obtenemos por fin la respuesta a la pregunta que se hizo anteriormente: *todo sistema de vectores linealmente independiente maximal del espacio vectorial de n dimensiones consta de n vectores.*

De los resultados obtenidos se pueden deducir también otras consecuencias.

Si en un sistema dado de vectores, linealmente dependiente, se han tomado dos subsistemas linealmente independientes maximales, o sea, dos subsistemas a los cuales no se les puede agregar otro vector del sistema sin violar la independencia lineal, entonces estos subsistemas contienen un número igual de vectores.

En efecto, si en el sistema de vectores

$$\alpha_1, \alpha_2, \ldots, \alpha_r \tag{13}$$

el subsistema

$$\alpha_1, \alpha_2, \ldots, \alpha_s, \ s < r, \tag{14}$$

es linealmente independiente maximal, entonces cualquiera de los vectores $\alpha_{s+1}, \ldots, \alpha_r$ se expresará linealmente mediante el sistema (14). Por otra parte, cualquier vector α_i del sistema (14) se expresa linealmente mediante este sistema: es suficiente tomar el mismo vector α_i con el coeficiente 1, y todos los demás vectores del sistema con el coeficiente 0. Ahora se ve fácilmente que los sistemas (13) y (14) son equivalentes. De aquí se deduce que el sistema (13) es equivalente a cualquiera de sus subsistemas linealmente independiente maximales, por consiguiente, todos estos subsistemas son equivalentes entre sí y, siendo linealmente independientes, contienen un mismo número de vectores.

El número de vectores de cualquier subsistema linealmente independiente maximal de un sistema dado de vectores, se llama *rango* de este sistema. Empleando esta noción, deduzcamos otra consecuencia más del teorema fundamental.

Sean dados dos sistemas de vectores de n dimensiones

$$\alpha_1, \alpha_2, \ldots, \alpha_r \tag{15}$$

y

$$\beta_1, \beta_2, \ldots, \beta_s, \tag{16}$$

no necesariamente linealmente independientes, y sea k, el rango del sistema (15) y l, el rango del sistema (16). Si el primer sistema se expresa linealmente mediante el segundo, entonces $k \leqslant l$. Si estos sistemas son equivalentes, $k = l$.

En efecto, sean

$$\alpha_{i_1}, \alpha_{i_2}, \ldots, \alpha_{i_k} \tag{17}$$

y

$$\beta_{j_1}, \beta_{j_2}, \ldots, \beta_{j_l} \tag{18}$$

subsistemas arbitrarios linealmente independientes maximales de los sistemas (15) y (16), respectivamente. Entonces, los sistemas (15) y (17) son equivalentes entre sí; esto mismo se refiere a los sistemas (16) y (18). Como el sistema (15) se expresa linealmente mediante el sistema (16), resulta ahora que el sistema (17) también se expresa linealmente mediante el sistema (16) y, por consiguiente, mediante el sistema (18), equivalente a él, después de lo cual no queda más que aplicar el teorema fundamental, empleando la independencia lineal del sistema (17). La segunda afirmación de la consecuencia que demostramos se deduce inmediatamente de la primera.

§ 10. Rango de una matriz

Dado un sistema de vectores de n dimensiones, surge la pregunta natural. ¿Es linealmente dependiente este sistema o no lo es? No se puede esperar que en cada caso concreto se obtenga sin dificultad la solución de este problema. Con un examen superficial sería difícil observar alguna dependencia lineal del sistema de vectores

$$\alpha = (2, -5, 1, -1), \ \beta = (1, 3, 6, 5), \ \gamma = (-1, 4, 1, 2),$$

a pesar de que, en realidad, estos vectores están ligados por la relación

$$7\alpha - 3\beta + 11\gamma = 0.$$

El § 1 proporciona un método para la resolución de este problema; como son conocidas las componentes de los vectores considerados, llamando incógnitas a los coeficientes de la dependencia lineal buscada, obtenemos un sistema de ecuaciones lineales homogéneas, que se resuelve por el método de Gauss. En el presente párrafo se indicará otro método para abordar el problema considerado; a la vez, nos aproximaremos considerablemente a nuestro objetivo principal, que consiste en resolver sistemas arbitrarios de ecuaciones lineales.

Sea dada la matriz

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{s1} & a_{s2} & \dots & a_{sn} \end{pmatrix},$$

con s filas y n columnas, donde los números s y n no están ligados de ningún modo. Las columnas de esta matriz, consideradas como vectores de s dimensiones, pueden ser, en general, linealmente dependientes. El rango del sistema de columnas, o sea, el número máximo de columnas linealmente independientes de la matriz A (con mayor precisión: el número de columnas que abarca cualquier subsistema linealmente independiente maximal del sistema de columnas), se llama *rango* de esta matriz.

Se sobreentiende, que se podrían considerar de modo semejante las filas de la matriz A como vectores de n dimensiones. Resulta que el rango del sistema de filas de la matriz es igual al rango del sistema de sus columnas, es decir, es igual al rango de esta matriz. La demostración de esta inesperada afirmación se obtendrá después de que indiquemos otra forma más de definir el rango de la matriz, lo que proporcionará a la vez un método para su cálculo.

Generalicemos primero el concepto de menor al caso de matrices rectangulares. Elijamos arbitrariamente en la matriz A, k filas y k columnas, $k \leqslant \min(s, n)$. Los elementos situados en las intersecciones de estas filas y columnas forman una matriz cuadrada de k-ésimo orden, cuyo determinante se llama *menor de k-ésimo orden* de la matriz A. A continuación, nos van a interesar los órdenes de los menores de la matriz A, que son diferentes de cero, y, precisamente, el **mayor de estos órdenes**. Para hallarlo es conveniente tener en cuenta la siguiente observación: si todos *los menores de k-ésimo orden de la matriz A son iguales a cero, entonces también son iguales a cero todos los menores de orden superior*. En efecto, desarrollando cualquier menor de orden $k + j$, $k < k + j \leqslant \min(s, n)$, por los menores de cualesquiera k filas, representamos este menor, según al teorema de Laplace, en forma de una suma de menores de orden k, multiplicados por ciertos menores de orden j, con lo que se demuestra que el menor de orden $k + j$ es igual a cero.

Demostremos ahora el siguiente **teorema sobre el rango de una matriz:**

El orden superior de los menores, diferentes de cero, de una matriz A, es igual al rango de esta matriz.

Demostración. Sea r el orden superior de los menores de la matriz A, diferentes de cero. Supongamos —lo que no restringe la gene-

ralidad de la demostración—, que el menor D, de r-ésimo orden, situado en el ángulo superior de la izquierda de la matriz A

$$A = \begin{pmatrix} a_{11} \dots a_{1r} & a_{1,r+1} \dots a_{1n} \\ \dots D \dots & \dots\dots\dots \\ a_{r1} \dots a_{rr} & a_{r,r+1} \dots a_{rn} \\ a_{r+1,1} \dots a_{r+1,r} & a_{r+1,r+1} \dots a_{r+1,n} \\ \dots\dots\dots & \dots\dots\dots \\ a_{s1} \dots a_{sr} & a_{s,r+1} \dots a_{sn} \end{pmatrix},$$

es diferente de cero, $D \neq 0$. Entonces, las primeras r columnas de la matriz A serán linealmente independientes entre sí. Si hubiese alguna dependencia lineal entre éstas, entonces, como al sumar los vectores *se suman sus componentes*, entre las columnas del menor D existiría la misma dependencia lineal y, por consiguiente, el menor D sería igual a cero.

Demostremos ahora que *cualquier l-ésima columna de la matriz A, $r < l \leqslant n$, es combinación lineal de las primeras r columnas.* Tomemos cualquier i, $1 \leqslant i \leqslant s$, y formemos el determinante auxiliar de $(r+1)$-ésimo orden

$$\Delta_i = \begin{vmatrix} a_{11} \dots a_{1r} a_{1l} \\ \dots\dots\dots \\ a_{r1} \dots a_{rr} a_{rl} \\ a_{i1} \dots a_{ir} a_{il} \end{vmatrix},$$

que se obtiene «orlando» el menor D con los elementos correspondientes de la l-ésima columna y de la i-ésima fila. Para cualquier i, el determinante Δ_i es igual a cero. En efecto, si $i > r$, entonces Δ_i será un menor de $(r+1)$-ésimo orden de nuestra matriz A, y, por lo tanto, es igual a cero, en virtud de la elección del número r. Si $i \leqslant r$, entonces Δ_i no será ya un menor de la matriz A, puesto que no puede ser obtenido de esta matriz suprimiendo algunas de sus filas y columnas; sin embargo, el determinante Δ_i contendrá ahora dos filas iguales y, *por consiguiente*, será de nuevo igual a cero.

Consideremos los complementos algebraicos de los elementos de la última fila del determinante Δ_i. Es evidente que el menor D sirve de complemento algebraico para el elemento a_{il}. Si $1 \leqslant j \leqslant r$, el complemento algebraico del elemento a_{ij} en Δ_i será el número

$$A_j = (-1)^{(r+1)+j} \begin{vmatrix} a_{11} \dots a_{1,j-1} a_{1,j+1} \dots a_{1r} a_{1l} \\ \dots\dots\dots\dots\dots \\ a_{r1} \dots a_{r,j-1} a_{r,j+1} \dots a_{rr} a_{rl} \end{vmatrix};$$

éste no depende de i y por eso se ha designado con A_j. Por lo tanto, desarrollando el determinante Δ_i por los elementos de su última fila e igualando a cero este desarrollo, puesto que $\Delta_i = 0$, obtenemos:

$$a_{i1}A_1 + a_{i2}A_2 + \ldots + a_{ir}A_r + a_{il}D = 0,$$

de donde, en virtud de que $D \neq 0$,

$$a_{il} = -\frac{A_1}{D} a_{i1} - \frac{A_2}{D} a_{i2} - \ldots - \frac{A_r}{D} a_{ir}.$$

Esta igualdad se verifica para todos los i, $i = 1, 2, \ldots, s$ y como sus coeficientes no dependen de i, resulta que toda la l-ésima columna de la matriz A es una suma de sus primeras r columnas, tomadas respectivamente con los coeficientes $-\frac{A_1}{D}, -\frac{A_2}{D}, \ldots, \frac{A_r}{D}$.

Por lo tanto, en el sistema de las columnas de la matriz A hemos hallado un subsistema linealmente independiente maximal compuesto de r columnas. Con esto queda demostrado que el rango de la matriz A es igual a r, es decir, queda demostrado el teorema sobre el rango.

Este teorema proporciona un método para el cálculo práctico del rango de la matriz, y también para la solución del problema sobre la existencia de dependencia lineal en un sistema dado de vectores; formando una matriz para la que los vectores dados sirvan de columnas, y calculando el rango de esta matriz, obtenemos el número mayor de vectores de nuestro sistema, linealmente independientes.

El método para el cálculo del rango de una matriz, basado en el teorema sobre el rango, requiere el cálculo de un número de menores de esta matriz que, aunque es finito, puede ser muy grande. Sin embargo, la siguiente observación da la posibilidad de introducir en este método simplificaciones considerables. Si el lector examina otra vez más la demostración del teorema sobre el rango de la matriz, observará que al efectuarla no se aplicó la igualdad a cero de *todos* los menores de $(r + 1)$-ésimo orden de la matriz A, sino que se usaron solamente los menores de $(r + 1)$-ésimo orden que orlaban al menor dado D de r-ésimo orden, diferente de cero (o sea, que lo contienen totalmente dentro de sí). Por lo tanto, de la igualdad a cero solamente de estos menores, se deduce que r es el máximo número de columnas linealmente independientes de la matriz A. Esto último trae consigo la igualdad a cero de todos los menores de $(r + 1)$-ésimo orden de esta matriz. Llegamos a la siguiente **regla para el cálculo del rango de una matriz:**

Al calcular el rango de una matriz se debe pasar de los menores de menor orden a los de orden mayor. Habiendo hallado un menor D de k-ésimo orden diferente de cero, se deben calcular solamente

los menores de $(k+1)$-ésimo orden que orlan al menor D: si todos éstos son iguales a cero, el rango de esta matriz es igual a k.

Ejemplos.

1. Hallar el rango de la matriz

$$A=\begin{pmatrix} 2 & -4 & 3 & 1 & 0 \\ 1 & -2 & 1 & -4 & 2 \\ 0 & 1 & -1 & 3 & 1 \\ 4 & -7 & 4 & -4 & 5 \end{pmatrix}.$$

El menor de segundo orden, situado en el ángulo superior de la izquierda de esta matriz, es igual a cero. Sin embargo, en esta matriz hay también menores de segundo orden, diferentes de cero, por ejemplo,

$$d=\begin{vmatrix} -4 & 3 \\ -2 & 1 \end{vmatrix} \neq 0.$$

El menor de tercer orden

$$d'=\begin{vmatrix} 2 & -4 & 3 \\ 1 & -2 & 1 \\ 0 & 1 & -1 \end{vmatrix},$$

es un orlado del menor d, diferente de cero, $d'=1$, no obstante, los orlados de cuarto orden del menor d' son iguales a cero:

$$\begin{vmatrix} 2 & -4 & 3 & 1 \\ 1 & -2 & 1 & -4 \\ 0 & 1 & -1 & 3 \\ 4 & -7 & 4 & -4 \end{vmatrix}=0, \quad \begin{vmatrix} 2 & -4 & 3 & 0 \\ 1 & -2 & 1 & 2 \\ 0 & 1 & -1 & 1 \\ 4 & -7 & 4 & 5 \end{vmatrix}=0.$$

Por lo tanto, el rango de la matriz A es igual a tres.

2. Hallar un subsistema, linealmente independiente, maximal en el sistema de vectores

$$\alpha_1=(2,\ -2,\ -4),\ \alpha_2=(1,\ 9,\ 3),\ \alpha_3=(-2,\ -4,\ 1),\ \alpha_4=(3,\ 7,\ -1).$$

Formamos la matriz

$$\begin{pmatrix} 2 & 1 & -2 & 3 \\ -2 & 9 & -4 & 7 \\ -4 & 3 & 1 & -1 \end{pmatrix},$$

en la que los vectores dados sirven de columnas. El rango de esta matriz es igual a dos: el menor de segundo orden situado en el ánguo superior de la izquierda es diferente de cero, pero los dos menores orlados de él, de tercer orden, son iguales a cero. De aquí se deduce que los vectores α_1, α_2 forman en el sistema dado uno de los subsistemas linealmente independientes maximales.

Como consecuencia del teorema sobre el rango de una matriz, demostremos la afirmación ya enunciada anteriormente:

El máximo número de filas linealmente independientes de cualquier matriz es igual al máximo número de sus columnas linealmente independientes, es decir, es igual al rango de la matriz.

Para la demostración, trasponemos la matriz, o sea, sus filas las hacemos columnas, conservando su numeración. En la transposición, el máximo orden de los menores de la matriz diferentes de cero no puede alterarse, puesto que la trasposición no altera al determinante y para cualquier menor de la matriz inicial, el menor obtenido de él por transposición está contenido en la nueva matriz y viceversa. De aquí se deduce que el rango de la nueva matriz es igual al rango de la matriz inicial; éste a su vez es igual al máximo número de columnas linealmente independientes de la nueva matriz, es decir, igual al máximo número de filas linealmente independientes de la matriz inicial.

Ejemplo. En el § 8 se introdujo el concepto de la forma lineal en n incógnitas y se dio la definición de suma de formas lineales y de su producto por un número. Esta definición permite generalizar el concepto de dependencia lineal, con todas sus propiedades, para el caso de formas lineales.

Sea dado el sistema de formas lineales

$$f_1 = x_1 + 2x_2 + x_3 + 3x_4,$$
$$f_2 = 4x_1 - x_2 + 5x_3 - 6x_4,$$
$$f_3 = x_1 - 3x_2 - 4x_3 - 7x_4,$$
$$f_4 = 2x_1 - x_2 - x_3.$$

Se necesita elegir en él un subsistema linealmente independiente maximal.

Formemos la matriz de los coeficientes de estas formas:

$$\begin{pmatrix} 1 & 2 & 1 & 3 \\ 4 & -1 & -5 & -6 \\ 1 & -3 & -4 & -7 \\ 2 & 1 & -1 & 0 \end{pmatrix}$$

y hallemos su rango. El menor de segundo orden, situado en el ángulo superior de la izquierda, es diferente de cero. Pero, como fácilmente se comprueba, sus cuatro determinantes orlados de tercer orden son iguales a cero. De aquí se deduce que las primeras dos filas de nuestra matriz son linealmente independientes, mientras que la tercera y la cuarta son combinaciones lineales de ellas. Por consiguiente, el sistema f_1, f_2 es el subsistema buscado del sistema dado de formas lineales.

Señalemos otra consecuencia importante del teorema sobre el rango de una matriz.

Un determinante de n-ésimo orden es igual a cero cuando, y sólo cuando, entre sus filas existe una dependencia lineal.

En una dirección, esta afirmación ya está demostrada en el § 4 (propiedad 8). Supongamos ahora que se ha dado un determinante de n-ésimo orden igual a cero o, en otras palabras, una matriz cuadrada de n-ésimo orden, cuyo único menor de máximo orden es igual a cero. De aquí se deduce que el máximo orden de los menores

de la matriz que son diferentes de cero es menor que n, o sea, que el rango es menor que n, y, por lo demostrado anteriormente, las filas de esta matriz son linealmente dependientes.

Se sobreentiende que en el enunciado de la consecuencia que hemos demostrado se puede hablar de las columnas del determinante, en lugar de las filas.

Existe también otro método para calcular el rango de una matriz que no está ligado con el teorema sobre el rango y que no requiere el cálculo de determinantes. Pero se puede aplicar solamente cuando se quiera determinar el mismo rango y no interese saber qué columnas (o filas) son las que precisamente forman un sistema linealmente independiente maximal. Veamos este método.

Se llaman *transformaciones elementales* de una matriz A a las siguientes:

(a) la permutación (trasposición) de dos filas o de dos columnas;

(b) la multiplicación de una fila (o de una columna) por un número arbitrario diferente de cero;

(c) la suma a una fila (o a una columna) de otra fila (columna) multiplicada por un número.

Fácilmente se observa que las *transformaciones elementales no alteran el rango de la matriz*. En efecto, si, por ejemplo, se aplican estas transformaciones a las columnas de la matriz, entonces el sistema de columnas, consideradas como vectores, se sustituye por otro equivalente. Demostremos esto solamente para la transformación (c), puesto que para las (a) y (b), es evidente. Supongamos que a la i-ésima columna se agrega la j-ésima columna, multiplicada por el número k. Si antes de la transformación, los vectores

$$\alpha_1, \ldots, \alpha_i, \ldots, \alpha_j, \ldots, \alpha_n, \tag{1}$$

servían de columnas de la matriz, después de la transformación servirán de columnas los vectores

$$\alpha_1, \ldots, \alpha_i' = \alpha_i + k\alpha_j, \ldots, \alpha_j, \ldots, \alpha_n. \tag{2}$$

El sistema (2) se expresa linealmente mediante el sistema (1). La igualdad

$$\alpha_i = \alpha_i' - k\alpha_j$$

muestra a su vez, que el sistema (1) se expresa linealmente mediante el (2). Por consiguiente, estos sistemas son equivalentes, y sus subsistemas, linealmente independientes maximales están compuestos de un mismo número de vectores.

Por lo tanto, para calcular el rango de una matriz, se puede simplificar previamente mediante una combinación de transformaciones elementales.

Se dice que una matriz que consta de s filas y n columnas es de *forma diagonal*, si todos sus elementos son iguales a cero, a excepción de los elementos a_{11}, a_{22}, ..., a_{rr} (donde $0 \leqslant r \leqslant \min(s, n)$), que son iguales a la unidad. Es evidente que el rango de esta matriz es igual a r.

Toda matriz se puede reducir a la forma diagonal mediante transformaciones elementales.

En efecto, sea dada la matriz

$$A = \begin{pmatrix} a_{11} & \ldots & a_{1n} \\ \cdot & \cdot & \cdot \\ a_{s1} & \ldots & a_{sn} \end{pmatrix}.$$

Si todos sus elementos son iguales a cero, esta tiene la forma diagonal. Si en ella hay elementos diferentes de cero, trasponiendo filas y columnas se puede conseguir que el elemento a_{11} sea diferente de cero. Multiplicando después la primera fila por a_{11}^{-1}, convertimos el elemento a_{11} en la unidad. Restando ahora de la j-ésima columna, $j > 1$, la primera columna, multiplicada por a_{1j}, se sustituye por cero el elemento a_{1j}. Efectuando esta transformación con todas las columnas, comenzando con la segunda, y también con todas las filas, llegaremos a la matriz de la forma:

$$A' = \begin{pmatrix} 1 & 0 & & 0 \\ 0 & a'_{22} & \dots & a'_{2n} \\ \cdot & \cdot & \cdot & \cdot \\ 0 & a'_{s2} & \dots & a'_{sn} \end{pmatrix}.$$

Luego efectuamos estas mismas transformaciones con la matriz que queda en el ángulo inferior de la derecha, etc., etc., Después de reiterar este proceso una cantidad finita de veces, llegaremos a la matriz diagonal que tiene el mismo rango que la matriz inicial A.

Por lo tanto, *para hallar el rango de una matriz hay que reducir esta matriz mediante transformaciones elementales a la forma diagonal y calcular el número de unidades que hay en su diagonal principal.*

Ejemplo. Hallar el rango de la matriz

$$A = \begin{pmatrix} 0 & 2 & -4 \\ -1 & -4 & 5 \\ 3 & 1 & 7 \\ 0 & 5 & -10 \\ 2 & 3 & 0 \end{pmatrix}.$$

Transponiendo en esta matriz la primera y segunda columna, y multiplicando la primera fila por el número $\frac{1}{2}$, llegamos a la matriz

$$\begin{pmatrix} 1 & 0 & -2 \\ -4 & -1 & 5 \\ 1 & 3 & 7 \\ 5 & 0 & -10 \\ 3 & 2 & 0 \end{pmatrix}.$$

Agregando a su tercera columna la primera duplicada y agregando después a cada una de las demás filas un múltiplo de la nueva primera fila, obtenemos la matriz

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & -3 \\ 0 & 3 & 9 \\ 0 & 0 & 0 \\ 0 & 2 & 6 \end{pmatrix}.$$

Finalmente, multiplicando la segunda fila por -1, restando de la tercera columna la segunda, multiplicada por tres, y restando de la tercera y quinta filas unos múltiplos de la segunda fila nueva, llegaremos a la forma diagonal buscada

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

Por lo tanto, el rango de la matriz A es igual a dos.

En el cap. 13 nos encontraremos otra vez con las transformaciones elementales y con la forma diagonal de la matriz; pero serán matrices cuyos elementos no serán números sino polinomios.

§ 11. Sistemas de ecuaciones lineales

Estudiaremos ahora sistemas arbitrarios de ecuaciones lineales, sin suponer que el número de ecuaciones sea igual al número de incógnitas. Nuestros resultados se podrán aplicar también al caso (que quedó sin examinar en el § 7) en el que el número de ecuaciones sea igual al número de incógnitas, siendo el determinante igual a cero.

Sea dado un sistema de ecuaciones lineales

$$\left.\begin{array}{l} a_{11}x_1+a_{12}x_2+\ldots+a_{1n}x_n=b_1, \\ a_{21}x_1+a_{22}x_2+\ldots+a_{2n}x_n=b_2, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ a_{s1}x_1+a_{s2}x_2+\ldots+a_{sn}x_n=b_s. \end{array}\right\} \qquad (1)$$

Como sabemos por el § 1, ante todo se debe resolver el problema sobre la compatibilidad de este sistema. Con este fin, tomemos la matriz A de los coeficientes del sistema y la matriz «ampliada» $\overline{A}$, obtenida al agregar a la matriz A la columna de los términos independientes,

$$A=\begin{pmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \ldots & \ldots & \ldots & \ldots \\ a_{s1} & a_{s2} & \ldots & a_{sn} \end{pmatrix}, \quad \overline{A}=\begin{pmatrix} a_{11} & a_{12} & \ldots & a_{1n} & b_1 \\ a_{21} & a_{22} & \ldots & a_{2n} & b_2 \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ a_{s1} & a_{s2} & \ldots & a_{sn} & b_s \end{pmatrix},$$

y calculemos los rangos de estas matrices. Es fácil ver que *el rango de la matriz $\overline{A}$, o es igual al rango de la matriz A, o es mayor en una unidad.* En efecto, tomemos un sistema máximal de columnas de la matriz A, linealmente independiente. Este también será linealmente independiente en la matriz $\overline{A}$. Si conserva también la propiedad de ser maximal, o sea, que la columna de los términos inde-

pendientes se expresa linealmente mediante el mismo, entonces los rangos de las matrices A y $\overline{A}$ son iguales; en caso contrario, agregando a este sistema la columna de los términos independientes, obtenemos el sistema linealmente independiente de columnas de la matriz $\overline{A}$, que en ésta será maximal.

El problema de la compatibilidad de un sistema de ecuaciones lineales se resuelve difinitivamente con el siguiente teorema.

Teorema de Kronecker-Capelli. *El sistema de ecuaciones* (1) *es compatible cuando, y sólo cuando, el rango de la matriz ampliada $\overline{A}$ es igual al rango de la matriz A.*

Demostración. 1. Supongamos que el sistema (1) es compatible y que $k_1, k_2, \ldots, k_n$ es una de sus soluciones. Sustituyendo estos números en lugar de las incógnitas del sistema (1), obtenemos s identidades, que muestran que la última columna de la matriz $\overline{A}$ es una suma de todas las demás columnas, tomadas con los coeficientes $k_1, k_2, \ldots, k_n$, respectivamente. Cualquiera otra columna de la matriz $\overline{A}$ forma parte también de la matriz A y, por eso, se expresa linealmente mediante todas las columnas de esta matriz. Recíprocamente, toda columna de la matriz A es también columna de la matriz $\overline{A}$, o sea, se expresa linealmente mediante las columnas de esta matriz. De aquí se deduce que los sistemas de columnas de las matrices A y $\overline{A}$ son equivalentes entre sí. Por consiguiente, como se demostró al final del § 9, estos dos sistemas de vectores de s dimensiones tienen un mismo rango; en otras palabras, los rangos de las matrices A y $\overline{A}$ son iguales entre sí.

2. Supongamos ahora que las matrices A y $\overline{A}$ tienen un mismo rango. De esto se deduce, que cualquier sistema linealmente independiente maximal de columnas de la matriz A se mantiene también en la matriz $\overline{A}$ como sistema linealmente independiente maximal. Por lo tanto, la última columna de la matriz $\overline{A}$ se expresa linealmente mediante este sistema y, por consiguiente, mediante el sistema de columnas de la matriz A. Así que existe un sistema de coeficientes $k_1, k_2, \ldots, k_n$ tal que la suma de las columnas de la matriz A, tomadas con estos coeficientes, es igual a la columna de los términos independientes. De aquí que los números $k_1, k_2, \ldots, k_n$ formen una solución del sistema (1). Por ello, la coincidencia de los rangos de las matrices A y $\overline{A}$ trae consigo la compatibilidad del sistema (1).

El teorema queda demostrado.

Al aplicar este teorema en los ejercicios prácticos, es necesario calcular primero el rango de la matriz A. Para esto hay que hallar uno de los menores de la matriz que sea diferente de cero y cuyos

orlados sean iguales a cero; sea éste el menor M. Después se deben calcular todos los menores de la matriz $\overline{A}$ que son orlados de M, pero que no estén contenidos en A (los llamados *determinantes característicos* del sistema (1)). Si todos éstos son iguales a cero, el rango de la matriz $\overline{A}$ es igual al rango de la matriz A y, por consiguiente, el sistema (1) es compatible; en caso contrario, es incompatible. De aquí que el teorema de Kronecker-Capelli se pueda enunciar del modo siguiente: el *sistema de ecuaciones lineales (1) es compatible cuando, y sólo cuando, todos sus determinantes característicos son iguales a cero.*

Supongamos ahora que el **sistema (1) es compatible.** El teorema de Kronecker—Capelli, mediante el que establecemos la compatibilidad de este sistema, afirma la existencia de una solución; mas éste no proporciona ningún método para la **averiguación práctica de todas las soluciones** del sistema. Pasemos ahora a resolver este problema.

Supongamos que la matriz A es de rango r. Como se ha demostrado en el párrafo anterior, r es el máximo número de filas linealmente independientes de la matriz A. Para precisar, supongamos que las primeras r filas de la matriz A son linealmente independientes, y que cada una de las demás es combinación lineal de ellas. Entonces, las primeras r filas de la matriz $\overline{A}$ serán también linealmente independientes: toda dependencia lineal entre ellas sería también una dependencia lineal entre las primeras r filas de la matriz A (¡véase la definición de la suma de vectores!). De la coincidencia de los rangos de las matrices A y $\overline{A}$ se deduce que las primeras r filas de la matriz $\overline{A}$ forman en ésta un sistema maximal de filas linealmente independiente, o sea, que cualquier otra fila de esta matriz es combinación lineal de ellas.

De aquí se deduce que cualquier ecuación del sistema (1), se puede representar como una suma de las primeras r ecuaciones, tomadas con ciertos coeficientes. En consecuencia, cualquier solución simultánea de las primeras r ecuaciones satisface también a todas las ecuaciones del sistema (1). Por consiguiente, es suficiente hallar todas las soluciones del sistema

$$\left.\begin{array}{l} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n = b_1, \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n = b_2, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ a_{r1}x_1 + a_{r2}x_2 + \ldots + a_{rn}x_n = b_r. \end{array}\right\} \quad (2)$$

Como las filas formadas por los coeficientes de las incógnitas en las ecuaciones (2) son linealmente independientes, o sea, la matriz de los coeficientes es de rango r, se tiene que $r \leqslant n$. Además,

al menos uno de los menores de r-ésimo orden de esta matriz es diferente de cero. Si $r = n$, entonces (2) será un sistema con igual número de ecuaciones que de incógnitas y con un determinante diferente de cero, por lo que éste y también el sistema (1), tendrán solución única, que es precisamente la que se calcula por la regla de Cramer.

Supongamos ahora que $r < n$, y para precisar, supongamos que es diferente de cero el menor de r-ésimo orden, formado por los coeficientes de las primeras r incógnitas. Traslademos al segundo miembro, en cada una de las ecuaciones (2), todos los términos que contienen las incógnitas $x_{r+1}, \ldots, x_n$ y elijamos para estas incógnitas algunos valores $c_{r+1}, \ldots, c_n$. Obtenemos un sistema de r ecuaciones

$$\left.\begin{array}{l} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1r}x_r = b_1 - a_{1,\,r+1}c_{r+1} - \ldots - a_{1n}c_n, \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2r}x_r = b_2 - a_{2,\,r+1}c_{r+1} - \ldots - a_{2n}c_n, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ a_{r1}x_1 + a_{r2}x_2 + \ldots + a_{rr}x_r = b_r - a_{r,\,r+1}c_{r+1} - \ldots - a_{rn}c_n \end{array}\right\} \quad (3)$$

con respecto a r incógnitas, $x_1, x_2, \ldots, x_r$. A este sistema se le puede aplicar la regla de Cramer, poseyendo por lo tanto, una solución única, $c_1, c_2, \ldots, c_r$; es evidente que el sistema de números $c_1, c_2, \ldots, c_r, c_{r+1}, \ldots, c_n$ representa una solución del sistema (2). Como los valores $c_{r+1}, \ldots, c_n$, para las incógnitas $x_{r+1}, \ldots, x_n$, llamadas *incógnitas independientes*, podían ser elegidos arbitrariamente, se pueden obtener de este modo infinitas soluciones distintas del sistema (2).

Por otra parte, toda solución del sistema (2) se puede obtener por el método indicado: si se ha obtenido alguna solución $c_1, c_2, \ldots, c_n$ del sistema (2), como valores para las incógnitas independientes tomamos los números $c_{r+1}, \ldots, c_n$. Entonces, los números $c_1, c_2, \ldots, c_r$ serán solución del sistema (3) y, por consiguiente, formarán la única solución de este sistema, que se calcula por la regla de Cramer.

Todo lo expuesto anteriormente se resume en la siguiente **regla para la solución de un sistema arbitrario de ecuaciones lineales:**

Sea dado un sistema compatible de ecuaciones lineales (1) *y sea r el rango de la matriz A de los coeficientes del sistema. Elijamos en A, r filas linealmente independientes y dejemos en el sistema* (1) *solamente aquellas ecuaciones, cuyos coeficientes forman parte de las filas elegidas. Dejemos en los primeros miembros de estas ecuaciones r incógnitas, de modo que el determinante formado por los coeficientes de ellas sea diferente de cero, mientras que las otras incógnitas las consideramos independientes, trasladándolas a los segundos miembros de las ecuaciones. Dando valores numéricos arbitrarios a las incógnitas*

independientes y calculando los valores de las demás incógnitas por la regla de Cramer, obtenemos todas las soluciones del sistema (1).

He aquí de nuevo el enunciado del resultado obtenido:

Un sistema compatible (1) *tiene solución única cuando, y sólo cuando, el rango de la matriz A es igual al número de las incógnitas.*

Ejemplos. 1. Resolver el sistema

$$\left.\begin{aligned} 5x_1 - x_2 + 2x_3 + x_4 &= 7, \\ 2x_1 + x_2 + 4x_3 - 2x_4 &= 1, \\ x_1 - 3x_2 - 6x_3 + 5x_4 &= 0. \end{aligned}\right\}$$

El rango de la matriz de los coeficientes es igual a dos: el menor de segundo orden, situado en el ángulo superior de la izquierda de esta matriz, es diferente de cero, pero ambos menores orlados de tercer orden son iguales a cero. El rango de la matriz ampliada es igual a tres, puesto que

$$\begin{vmatrix} 5 & -1 & 7 \\ 2 & 1 & 1 \\ 1 & -3 & 0 \end{vmatrix} = -35 \neq 0.$$

De aquí se deduce que el sistema es incompatible.

2. Resolver el sistema

$$\left.\begin{aligned} 7x_1 + 3x_2 &= 2, \\ x_1 - 2x_2 &= -3, \\ 4x_1 + 9x_2 &= 11. \end{aligned}\right\}$$

El rango de la matriz de los coeficientes es igual a dos, o sea, es igual al número de incógnitas; el rango de la matriz ampliada también es igual a dos. Por lo tanto, el sistema es compatible y tiene solución única. Los primeros miembros de las primeras dos ecuaciones son linealmente independientes; resolviendo el sistema de estas dos ecuaciones, obtenemos los siguientes valores para las incógnitas:

$$x_1 = -\frac{5}{17}, \quad x_2 = \frac{23}{17}.$$

Vemos fácilmente que esta solución satisface también a la tercera ecuación.

3. Resolver el sistema

$$\left.\begin{aligned} x_1 + x_2 - 2x_3 - x_4 + x_5 &= 1, \\ 3x_1 - x_2 + x_3 + 4x_4 + 3x_5 &= 4, \\ x_1 + 5x_2 - 9x_3 - 8x_4 + x_5 &= 0. \end{aligned}\right\}$$

El sistema es compatible, puesto que el rango de la matriz ampliada al igual que el de la matriz de los coeficientes es igual a dos. Los primeros miembros de la primera y tercera ecuaciones son linealmente independientes, puesto que los coeficientes de las incógnitas x_1 y x_2 forman un menor de segundo orden diferente de cero. El sistema de estas dos ecuaciones lo resolvemos suponiendo que las incógnitas x_3, x_4, x_5 son independientes; para ello, trasladamos éstas a los segundos miembros de las ecuaciones y suponemos que ya se les han atribuido valores numéricos. Aplicando la regla de Cramer, obtenemos:

$$x_1 = \frac{5}{4} + \frac{1}{4}x_3 - \frac{3}{4}x_4 - x_5,$$

$$x_2 = -\frac{1}{4} + \frac{7}{4}x_3 + \frac{7}{4}x_4.$$

Estas igualdades determinan la *solución general* del sistema dado: asignando a las incógnitas independientes valores numéricos arbitrarios, obtenemos todas las soluciones de nuestro sistema. Así pues, son soluciones de nuestro sistema, por ejemplo, los vectores $(2, 5, 3, 0, 0)$, $(3, 5, 2, 1, -2)$, $(0, -\frac{1}{4}, -1, 1, \frac{1}{4})$, etc. Por otra parte, sustituyendo las expresiones para x_1 y x_2 de la solución general en cualquiera de las ecuaciones del sistema, por ejemplo, en la segunda, que fue anteriormente excluida, obtenemos una identidad.

4. Resolver el sistema

$$\left.\begin{aligned} 4x_1+x_2-2x_3+x_4&=3,\\ x_1-2x_2-x_3+2x_4&=2,\\ 2x_1+5x_2-x_4&=-1,\\ 3x_1+3x_2-x_3-3x_4&=1. \end{aligned}\right\}$$

A pesar de que el número de ecuaciones es igual al número de incógnitas, no se puede aplicar la regla de Cramer, pues el determinante del sistema es igual a cero. El rango de la matriz de los coeficientes es igual a tres: en el ángulo superior de la derecha de esta matriz está situado un menor de tercer orden, diferente de cero. El rango de la matriz ampliada también es igual a tres, es decir, el sistema es compatible. Examinando solamente las primeras tres ecuaciones y tomando la incógnita x_1 como independiente, obtenemos la solución general en la forma:

$$x_2=-\frac{1}{5}-\frac{2}{5}x_1,\quad x_3=-\frac{8}{5}+\frac{9}{5}x_1,\quad x_4=0.$$

5. Sea dado un sistema compuesto de $n+1$ ecuaciones respecto a n incógnitas. La matriz ampliada $\overline{A}$ de este sistema es cuadrada, de orden $n+1$. Si nuestro sistema es compatible, entonces, según el teorema de Kronecker-Capelli, el determinante de la matriz $\overline{A}$ tiene que ser igual a cero.

Así, pues, sea dado el sistema

$$\left.\begin{aligned} x_1-8x_2&=3,\\ 2x_1+x_2&=1,\\ 4x_1+7x_2&=-4. \end{aligned}\right\}$$

El determinante de los coeficientes y de los términos independientes de estas ecuaciones es diferente de cero:

$$\begin{vmatrix} 1 & -8 & 3\\ 2 & 1 & 1\\ 4 & 7 & -4 \end{vmatrix}=-77,$$

por lo tanto, el sistema es incompatible.

En general, la afirmación recíproca no es justa: de la igualdad a cero del determinante de la matriz $\overline{A}$ no se deduce la coincidencia de los rangos de las matrices A y $\overline{A}$.

§ 12. Sistemas de ecuaciones lineales homogéneas

Apliquemos los resultados del párrafo anterior al caso de un *sistema de ecuaciones lineales homogéneas*:

$$\left.\begin{array}{l} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n = 0, \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n = 0, \\ \cdots\cdots\cdots\cdots\cdots\cdots \\ a_{s1}x_1 + a_{s2}x_2 + \ldots + a_{sn}x_n = 0. \end{array}\right\} \qquad (1)$$

Del teorema de Kronecker-Capelli se deduce que este sistema siempre es compatible, puesto que, agregando una columna de ceros, no se puede elevar el rango de la matriz. Por cierto, esto se observa inmediatamente, ya que el sistema (1) siempre posee la *solución nula* $(0, 0, \ldots, 0)$.

Supongamos que la matriz A de los coeficientes del sistema (1) es de rango r. *Si $r = n$, la solución nula es la única solución del sistema* (1); *si $r < n$, el sistema posee también soluciones diferentes de la nula*; para hallar todas estas soluciones se emplea el mismo método que anteriormente se usó en el caso de un sistema arbitrario de ecuaciones. En particular, un *sistema de n ecuaciones lineales homogéneas con n incógnitas tiene soluciones diferentes de la nula cuando, y sólo cuando, el determinante de este sistema es igual a cero**. En efecto, la igualdad a cero de este determinante es equivalente a la afirmación de que el rango de la matriz A es menor que n. Por otra parte, *si en el sistema de ecuaciones lineales homogéneas el número de ecuaciones es menor que el número de incógnitas, el sistema posee indispensablemente soluciones diferentes de la nula*, puesto que en este caso el rango no puede ser igual al número de las incógnitas; este resultado fue obtenido en el § 1 por medio de otros razonamientos.

Veamos en particular el caso de un sistema, compuesto de $n-1$ ecuaciones homogéneas con respecto a n incógnitas, en el que se supone que los primeros miembros de las ecuaciones son linealmente independientes entre sí. Sea

$$A = \begin{pmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{n-1,1} & a_{n-1,2} & \ldots & a_{n-1,n} \end{pmatrix}$$

la matriz de los coeficientes de este sistema; designemos con M_i el menor de $(n-1)$-ésimo orden, obtenido después de suprimir en la matriz A la i-ésima columna, $i = 1, 2, \ldots, n$. *Entonces, una de las soluciones de nuestro sistema será el sistema de números*

$$M_1, \ -M_2, \ M_3, \ -M_4, \ \ldots, \ (-1)^{n-1} M_n, \qquad (2)$$

* Una mitad de esta afirmación se demostró ya en el § 7.

y cualquier otra solución será proporcional a ésta.

Demostración. Como, por hipótesis, el rango de la matriz A es igual a $n-1$, uno de los menores M_i tiene que ser diferente de cero; supongamos que sea M_n. Tomemos la incógnita x_n como independiente y la traslademos al segundo miembro en cada una de las ecuaciones, obteniendo:

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1,\,n-1}x_{n-1} &= -a_{1n}x_n, \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2,\,n-1}x_{n-1} &= -a_{2n}x_n, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ a_{n-1,\,1}x_1 + a_{n-1,\,2}x_2 + \ldots + a_{n-1,\,n-1}x_{n-1} &= -a_{n-1,\,n}x_n. \end{aligned}$$

Aplicando ahora la regla de Cramer, obtenemos la solución general del sistema dado de ecuaciones, la cual, después de sencillas transformaciones, puede ser expresada de la forma siguiente:

$$x_i = (-1)^{n-1}\frac{M_i}{M_n}x_n, \quad i = 1,\ 2,\ \ldots,\ n-1. \tag{3}$$

Haciendo $x_n = (-1)^{n-1} M_n$, obtenemos: $x_i = (-1)^{2n-i-1} M_i$, $i = 1, 2, \ldots, n-1$, o bien, como la diferencia $(2n-i-1)-(i-1) = 2n-2i$ es un número par, $x_i = (-1)^{i-1} M_i$, es decir, el sistema de números (2) es verdaderamente una solución de nuestro sistema de ecuaciones. Cualquier otra solución de este sistema se obtiene de las fórmulas (3) con otro valor númerico de la incógnita x_n, por lo que será proporcional a la solución (2). Se comprende que la afirmación considerada es justa también cuando $M_n = 0$, siendo diferente de cero uno de los menores M_i, $1 \leqslant i \leqslant n-1$.

Las soluciones de un sistema de ecuaciones lineales homogéneas poseen las siguientes propiedades. Si el vector $\beta = (b_1, b_2, \ldots, b_n)$ es una solución del sistema (1), entonces, para cualquier número k, el vector $k\beta = (kb_1, kb_2, \ldots, kb_n)$ también será solución de este sistema. Esto se comprueba inmediatamente por sustitución en cualquiera de las ecuaciones (1). Si, además, el vector $\gamma = (c_1, c_2, \ldots, c_n)$ es otra solución del sistema (1), el vector $\beta + \gamma = (b_1 + c_1, b_2 + c_2, \ldots, b_n + c_n)$ también será solución de este sistema:

$$\sum_{j=1}^{n} a_{ij}(b_j + c_j) = \sum_{j=1}^{n} a_{ij}b_j + \sum_{j=1}^{n} a_{ij}c_j = 0, \quad i = 1,\ 2,\ \ldots,\ s.$$

Por eso, en general, *cualquier combinación lineal de soluciones del sistema homogéneo* (1) *será también solución de este sistema.* Obsérvese que en el caso de un sistema *no homogéneo*, o sea, de un sistema de ecuaciones lineales, cuyos términos independientes no son todos iguales a cero, la afirmación correspondiente no se cumple: ni la suma de dos soluciones de un sistema de ecuaciones no homogéneas, ni el producto de una solución de este sistema por un número, será ya solución de este sistema.

Por el § 9 se sabe que cualquier sistema de vectores de n dimensiones, compuesto de más de n vectores, es linealmente dependiente. De aquí se deduce que entre las soluciones del sistema homogéneo (1) (como es sabido, éstas son vectores de n dimensiones) se pue-

de elegir un sistema *finito* linealmente independiente maximal de modo que sea maximal en el sentido de que cualquier otra solución del sistema (1) sea combinación lineal de las soluciones que forman parte de este sistema elegido. Todo sistema maximal de soluciones linealmente independientes del sistema de ecuaciones homogéneas (1), se llama *sistema fundamental de soluciones.*

Subrayemos otra vez más que *un vector de n dimensiones es solución del sistema* (1) *si, y sólo si, éste es combinación lineal de los vectores que forman el sistema fundamental dado.*

Se entiende que existirá un sistema fundamental solamente en el caso en que el sistema (1) tenga soluciones no nulas, o sea, *cuando el rango de su matriz de los coeficientes sea menor que* el número de las incógnitas. En tal caso, el sistema (1) puede tener muchos sistemas de soluciones fundamentales diversas. Sin embargo, todos estos sistemas serán equivalentes entre sí, puesto que cada vector de cada uno de estos sistemas se expresa linealmente mediante cualquier otro sistema. Por ello, los sistemas *constan de un mismo número de soluciones.*

Subsiste el siguiente teorema:

Si el rango r de la matriz de los coeficientes de un sistema de ecuaciones lineales homogéneas (1) *es menor que el número de las incógnitas n, entonces cualquier sistema fundamental de soluciones del sistema* (1) *consta de* $n-r$ *soluciones.*

Para la demostración, observemos que $n-r$ es el número de incógnitas independientes en el sistema (1); supongamos que las incógnitas independientes son: $x_{r+1}, x_{r+2}, \ldots, x_n$. Consideremos un determinante cualquiera d, de orden $n-r$ diferente de cero, que lo escribiremos de la forma siguiente:

$$d=\begin{vmatrix} c_{1,\,r+1}, & c_{1,\,r+2}, & \ldots, & c_{1n} \\ c_{2,\,r+1}, & c_{2,\,r+2}, & \ldots, & c_{2n} \\ \cdot & \cdot & \cdot & \cdot \\ c_{n-r,\,r+1}, & c_{n-r,\,r+2}, & \ldots, & c_{n-r,\,n} \end{vmatrix}.$$

Tomando los elementos de la i-ésima fila de este determinante, $1\leqslant i\leqslant n-r$, como valores para las incógnitas independientes, obtenemos como es sabido, unos valores unívocamente determinados para las incógnitas $x_1, x_2, \ldots, x_r$, o sea, llegaremos a una solución completamente determinada del sistema de ecuaciones (1); escribamos esta solución en forma de vector

$$\alpha_i=(c_{i1}, c_{i2}, \ldots, c_{ir}, c_{i,\,r+1}, c_{i,\,r+2}, \ldots, c_{in}).$$

El sistema obtenido de vectores $\alpha_1, \alpha_2, \ldots, \alpha_{n-r}$ representa un sistema fundamental de soluciones del sistema de ecuaciones (1). En efecto, este sistema de vectores es linealmente independiente,

puesto que la matriz formada por estos vectores como filas contiene un menor d, de orden $n-r$, diferente de cero. Por otra parte, sea

$$\beta=(b_1,\ b_2,\ \ldots,\ b_r,\ b_{r+1},\ b_{r+2},\ \ldots,\ b_n)$$

una solución arbitraria del sistema de ecuaciones (1). Demostremos que el vector β se expresa linealmente mediante los vectores $\alpha_1,\ \alpha_2,\ \ldots,\ \alpha_{n-r}$.

Designemos con α'_i, $i=1,\ 2,\ \ldots,\ n-r$, la i-ésima fila del determinante d, considerada como un vector de $(n-r)$ dimensiones y sea

$$\beta'=(b_{r+1},\ b_{r+2},\ \ldots,\ b_n).$$

Los vectores α'_i, $i=1,\ 2,\ \ldots,\ n-r$ son linealmente independientes, ya que $d\neq 0$. Sin embargo, el sistema de vectores de $(n-r)$ dimensiones

$$\alpha'_1,\ \alpha'_2,\ \ldots,\ \alpha'_{n-r},\ \beta'$$

es linealmente dependiente, debido a que en éste el número de vectores es mayor que las dimensiones de ellos. Por consiguiente, existen unos números $k_1,\ k_2,\ \ldots,\ k_{n-r}$ tales que

$$\beta'=k_1\alpha'_1+k_2\alpha'_2+\ldots+k_{n-r}\alpha'_{n-r}. \tag{4}$$

Examinemos ahora el vector de n dimensiones

$$\delta=k_1\alpha_1+k_1\alpha_2+\ldots+k_{n-r}\alpha_{n-r}-\beta.$$

El vector δ, siendo combinación lineal de las soluciones del sistema de ecuaciones homogéneas (1), representa también una solución del mismo. De la igualdad (4) se deduce que en la solución δ los valores para todas las incógnitas independientes son iguales a cero. No obstante, la única solución del sistema de ecuaciones (1) que resulta con los valores iguales a cero para las incógnitas independientes, es la solución nula. Por lo tanto, $\delta=0$, de donde

$$\beta=k_1\alpha_1+k_2\alpha_2+\ldots+k_{n-r}\alpha_{n-r}.$$

El teorema queda demostrado.

Obsérvese que la demostración expuesta nos permite afirmar que tomando por d todos los determinantes posibles de orden $n-r$, diferentes de cero, obtenemos todos los sistemas fundamentales de soluciones del sistema de ecuaciones homogéneas (1).

Ejemplo. Sea dado el sistema de ecuaciones lineales homogéneas

$$\left.\begin{aligned} 3x_1+x_2-8x_3+2x_4+x_5&=0,\\ 2x_1-2x_2-3x_3-7x_4+2x_5&=0,\\ x_1+11x_2-12x_3+34x_4-5x_5&=0,\\ x_1-5x_2+2x_3-16x_4+3x_5&=0. \end{aligned}\right\}$$

El rango de la matriz de los coeficientes es igual a dos; el número de incógnitas es igual a cinco. Por consiguiente, cualquier sistema fundamental de soluciones de este sistema de ecuaciones consta de tres soluciones. Resolvamos el sistema limitándonos a las dos primeras ecuaciones linealmente independientes y tomando como independientes las incógnitas x_3, x_4, x_5. Obtenemos la solución general de la forma:

$$x_1=\frac{19}{8}x_3+\frac{3}{8}x_4-\frac{1}{2}x_5,$$

$$x_2=\frac{7}{8}x_3-\frac{25}{8}x_4+\frac{1}{2}x_5.$$

Tomemos luego los siguientes tres vectores de tres dimensiones, linealmente independientes: (1, 0, 0), (0, 1, 0), (0, 0, 1). Sustituyendo las componentes de cada uno de ellos en la solución general, en calidad de valores para las incógnitas independientes, y, calculando los valores para x_1 y x_2, obtenemos el siguiente sistema fundamental de soluciones del sistema dado de ecuaciones:

$$\alpha_1=\left(\frac{19}{8},\ \frac{7}{8},\ 1, 0, 0\right),\quad \alpha_2=\left(\frac{3}{8},\ -\frac{25}{8},\ 0, 1, 0\right),$$

$$\alpha_3=\left(-\frac{1}{2},\ \frac{1}{2},\ 0, 0, 1\right).$$

Por último, consideremos la relación que existe entre las soluciones de los sistemas *homogéneos* y *no homogéneos*. Sea dado un sistema de ecuaciones lineales no homogéneas:

$$\left.\begin{array}{l} a_{11}x_1+a_{12}x_2+\ldots+a_{1n}x_n=b_1,\\ a_{21}x_1+a_{22}x_2+\ldots\quad a_{2n}x_n=b_2,\\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\\ a_{s1}x_1+a_{s2}x_2+\ldots+a_{sn}x_n=b_s. \end{array}\right\} \qquad (5)$$

El sistema de ecuaciones lineales homogéneas

$$\left.\begin{array}{l} a_{11}x_1+a_{12}x_2+\ldots+a_{1n}x_n=0,\\ a_{21}x_1+a_{22}x_2+\ldots+a_{2n}x_n=0,\\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\\ a_{s1}x_1+a_{s2}x_2+\ldots+a_{sn}x_n=0, \end{array}\right\} \qquad (6)$$

obtenido del sistema (5) al sustituir por ceros los términos independientes, se llama *sistema reducido*. Entre las soluciones de los sistemas (5) y (6) existe una notable relación, como lo muestran perfectamente los dos teoremas siguientes:

I. La suma de cualquier solución del sistema (5) *con cualquier solución del sistema reducido* (6) *será nuevamente solución del sistema* (5).

En efecto, sea $c_1, c_2, \ldots, c_n$ una solución del sistema (5) y $d_1, d_2, \ldots, d_n$, una solución del sistema (6). Tomemos cualquiera de las ecuaciones del sistema (5), por ejemplo la k-ésima, y susti-

tuyamos en ella, en lugar de las incógnitas, los números $c_1 + d_1$, $c_2 + d_2$, . . ., $c_n + d_n$. Resulta:

$$\sum_{j=1}^{n} a_{kj}(c_j + d_j) = \sum_{j=1}^{n} a_{kj}c_j + \sum_{j=1}^{n} a_{kj}d_j = b_k + 0 = b_k.$$

II. La diferencia de dos soluciones cualesquiera del sistema es solución del sistema reducido (6).

En efecto, sean $c_1, c_2, \ldots, c_n$ y $c_1', c_2', \ldots, c_n'$ dos soluciones del sistema (5). Tomemos cualquiera de las ecuaciones del sistema (6), por ejemplo la k-ésima, y sustituyamos en ella, en lugar de las incógnitas, los números

$$c_1 - c_1',\ c_2 - c_2',\ \ldots,\ c_n - c_n'.$$

Resulta:

$$\sum_{j=1}^{n} a_{kj}(c_j - c_j') = \sum_{j=1}^{n} a_{kj}c_j - \sum_{j=1}^{n} a_{kj}c_j' = b_k - b_k = 0.$$

De estos teoremas se deduce que, *hallando una solución del sistema de ecuaciones lineales no homogéneas* (5) *y sumándola con cada una de las soluciones del sistema reducido* (6), *obtenemos todas las soluciones del sistema* (5).

CAPITULO III

ALGEBRA DE LAS MATRICES

§ 13. Multiplicación de matrices

En los capítulos anteriores el concepto de matriz se había empleado como un instrumento auxiliar, esencial para el estudio de los *sistemas de ecuaciones lineales. Las numerosas y diversas aplicaciones* de este concepto contribuyeron a convertirlo en el objetivo de una amplia teoría particular que, en gran parte, sale fuera de los márgenes de nuestro curso. Ahora nos ocuparemos de los fundamentos de esta teoría que comienza definiendo de un modo original, pero bien fundamentado, dos operaciones algebraicas: la suma y la multiplicación, aplicables al conjunto de todas las matrices cuadradas *de un orden dado. Examinemos primero* la definición del producto de matrices; la suma de matrices la veremos en el § 15.

Por el curso de geometría analítica se sabe que al girar los ejes de un sistema rectangular de coordenadas en el plano, en un ángulo α, las coordenadas de los puntos se transforman según las fórmulas siguientes:

$$x = x' \cos\alpha - y' \operatorname{sen}\alpha,$$
$$y = x' \operatorname{sen}\alpha + y' \cos\alpha,$$

donde x, y son las coordenadas primitivas del punto, mientras que x', y' son sus coordenadas nuevas; por lo tanto, x e y se expresan linealmente mediante x' e y', con ciertos coeficientes numéricos. En diversas ocasiones también nos encontramos con la necesidad de efectuar una transformación de las indeterminadas (o de las variables) tal, que las indeterminadas primitivas queden expresadas *linealmente mediante* las *nuevas; ordinariamente*, esta sustitución de las indeterminadas se llama transformación lineal (o sustitución lineal). Por consiguiente, llegamos a la siguiente definición:

Se llama *transformación lineal de las indeterminadas* al paso del sistema de n indeterminadas $x_1, x_2, \ldots, x_n$ al sistema de n indeterminadas $y_1, y_2, \ldots, y_n$, de manera que las indeterminadas primitivas queden expresadas linealmente mediante las nuevas

con ciertos coeficientes numéricos:

$$\left.\begin{array}{l} x_1 = a_{11}y_1 + a_{12}y_2 + \ldots + a_{1n}y_n, \\ x_2 = a_{21}y_1 + a_{22}y_2 + \ldots + a_{2n}y_n, \\ \cdots\cdots\cdots\cdots\cdots\cdots \\ x_n = a_{n1}y_1 + a_{n2}y_2 + \ldots + a_{nn}y_n. \end{array}\right\} \qquad (1)$$

La transformación lineal (1) se determina completamente por la matriz de los coeficientes

$$A = \begin{pmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \cdot & \cdot & \cdot & \cdot \\ a_{n1} & a_{n2} & \ldots & a_{nn} \end{pmatrix},$$

puesto que dos transformaciones lineales con una misma matriz pueden diferenciarse entre sí solamente en las letras que designan las indeterminadas; sin embargo, nosotros supondremos que la elección de estas notaciones corre a nuestro cargo. Recíprocamente, dada una matriz arbitraria de n-ésimo orden, podemos escribir inmediatamente una transformación lineal, para la que esta matriz sirva de matriz de sus coeficientes. Por lo tanto, entre las transformaciones lineales de n indeterminadas y las matrices cuadradas de n-ésimo orden existe una correspondencia biunívoca, y por ello, a cualquier noción ligada con las transformaciones lineales y a cualquier propiedad de estas transformaciones tiene que corresponder una noción o una propiedad análoga, referente a las matrices.

Examinemos la cuestión sobre la realización consecutiva de dos transformaciones lineales. Supongamos que después de la transformación lineal (1), se ha realizado la transformación lineal

$$\left.\begin{array}{l} y_1 = b_{11}z_1 + b_{12}z_2 + \ldots + b_{1n}z_n, \\ y_2 = b_{21}z_1 + b_{22}z_2 + \ldots + b_{2n}z_n, \\ \cdots\cdots\cdots\cdots\cdots\cdots \\ y_n = b_{n1}z_1 + b_{n2}z_2 + \ldots + b_{nn}z_n, \end{array}\right\} \qquad (2)$$

que sustituye al sistema de indeterminadas $y_1, y_2, \ldots, y_n$ por el sistema $z_1, z_2, \ldots, z_n$; designemos con B la matriz de esta transformación. Sustituyendo en (1) las expresiones para $y_1, y_2, \ldots, y_n$, dadas en (2), llegaremos a unas expresiones lineales para las indeterminadas $x_1, x_2, \ldots, x_n$ mediante las indeterminadas $z_1, z_2, \ldots, z_n$. Por lo tanto, el *resultado de la realización consecutiva de dos transformaciones lineales de las indeterminadas, es de nuevo una transformación lineal.*

Ejemplo. El resultado de la realización consesutiva de las transformaciones lineales

$$x_1 = 3y_1 - y_2, \qquad y_1 = z_1 + z_2,$$
$$x_2 = y_1 + 5y_2, \qquad y_2 = 4z_1 + 2z_2$$

es la transformación lineal

$$x_1 = 3(z_1 + z_2) - (4z_1 + 2z_2) = -z_1 + z_2,$$
$$x_2 = (z_1 + z_2) + 5(4z_1 + 2z_2) = 21z_1 + 11z_2.$$

Designemos con C la matriz de la transformación lineal que representa el resultado de la realización consecutiva de las transformaciones (1) y (2), y hallemos la ley por la que se expresan los elementos c_{ik}, $i, k = 1, 2, \ldots, n$ mediante los elementos de las matrices A y B. Escribiendo abreviadamente las transformaciones (1) y (2) en la forma

$$x_i = \sum_{j=1}^{n} a_{ij} y_j \ \ i = 1, 2, \ldots, n; \ y_j = \sum_{k=1}^{n} b_{jk} z_k, \ j = 1, 2, \ldots, n,$$

obtenemos

$$x_i = \sum_{j=1}^{n} a_{ij} \left(\sum_{k=1}^{n} b_{jk} z_k \right) = \sum_{k=1}^{n} \left(\sum_{j=1}^{n} a_{ij} b_{jk} \right) z_k, \ i = 1, 2, \ldots, n.$$

En consecuencia, el coeficiente de z_k en la expresión para x_i, es decir, el elemento c_{ik} de la matriz C, tiene la forma

$$c_{ik} = \sum_{j=1}^{n} a_{ij} b_{jk} = a_{i1} b_{1k} + a_{i2} b_{2k} + \ldots + a_{in} b_{nk}; \tag{3}$$

el elemento de la matriz C, situado en la i-ésima fila y en la k-ésima columna, es igual a la suma de los productos de los correspondientes elementos de la i-ésima fila de la matriz A y de la k-ésima columna de la matriz B.

La fórmula (3), que da la expresión de los elementos de la matriz C mediante los elementos de las matrices A y B, permite escribir inmediatamente la matriz C, siendo dadas las matrices A y B, sin recurrir a las transformaciones lineales correspondientes a estas matrices. De este modo, a cualquier par de matrices cuadradas de n-ésimo orden se pone en correspondencia una tercera matriz unívocamente determinada. Se puede decir que hemos definido una operación algebraica en el conjunto de todas las matrices cuadradas de n-ésimo orden; ésta se llama *multiplicación de las matrices*, y la matriz C, *producto* de la matriz A por la matriz B:

$$C = AB.$$

Enunciemos una vez más la relación entre las transformaciones lineales y el producto de las matrices:

La transformación lineal de las indeterminadas obtenida como resultado de la realización consecutiva de dos transformaciones lineales con las matrices A y B, tiene a la matriz AB por matriz de sus coeficientes.

Ejemplos

1) $$\begin{pmatrix} 4 & 9 \\ -1 & 3 \end{pmatrix} \cdot \begin{pmatrix} 1 & -3 \\ -2 & 1 \end{pmatrix} = \begin{pmatrix} 4\cdot 1+9\cdot(-2) & 4\cdot(-3)+9\cdot 1 \\ (-1)\cdot 1+3\cdot(-2) & (-1)\cdot(-3)+3\cdot 1 \end{pmatrix} = \begin{pmatrix} -14 & -3 \\ -7 & 6 \end{pmatrix}.$$

2) $$\begin{pmatrix} 2 & 0 & 1 \\ -2 & 3 & 2 \\ 4 & -1 & 5 \end{pmatrix} \cdot \begin{pmatrix} -3 & 1 & 0 \\ 0 & 2 & 1 \\ 0 & -1 & 3 \end{pmatrix} = \begin{pmatrix} -6 & 1 & 3 \\ 6 & 2 & 9 \\ -12 & -3 & 14 \end{pmatrix}.$$

3) $$\begin{pmatrix} 7 & 2 \\ 1 & 1 \end{pmatrix}^2 = \begin{pmatrix} 7 & 2 \\ 1 & 1 \end{pmatrix} \cdot \begin{pmatrix} 7 & 2 \\ 1 & 1 \end{pmatrix} = \begin{pmatrix} 51 & 16 \\ 8 & 3 \end{pmatrix}.$$

4) Hallar el resultado de la realización consecutiva de las transformaciones lineales

$$\begin{aligned} x_1 &= 5y_1 - y_2 + 3y_3, \\ x_2 &= y_1 - 2y_2, \\ x_3 &= 7y_2 - y_3 \end{aligned}$$

y

$$\begin{aligned} y_1 &= 2z_1 + z_3, \\ y_2 &= z_2 - 5z_3, \\ y_3 &= 2z_2. \end{aligned}$$

Multiplicando las matrices, obtenemos:

$$\begin{pmatrix} 5 & -1 & 3 \\ 1 & -2 & 0 \\ 0 & 7 & -1 \end{pmatrix} \cdot \begin{pmatrix} 2 & 0 & 1 \\ 0 & 1 & -5 \\ 0 & 2 & 0 \end{pmatrix} = \begin{pmatrix} 10 & 5 & 10 \\ 2 & -2 & 11 \\ 0 & 5 & -35 \end{pmatrix},$$

de donde, la transformación lineal buscada tiene la forma:

$$\begin{aligned} x_1 &= 10z_1 + 5z_2 + 10z_3, \\ x_2 &= 2z_1 - 2z_2 + 11z_3, \\ x_3 &= 5z_2 - 35z_3. \end{aligned}$$

Tomemos uno de los ejemplos que acabamos de estudiar de multiplicación de las matrices, por ejemplo el 2), y hallemos el producto de las mismas matrices, pero tomadas en orden inverso:

$$\begin{pmatrix} -3 & 1 & 0 \\ 0 & 2 & 1 \\ 0 & -1 & 3 \end{pmatrix} \cdot \begin{pmatrix} 2 & 0 & 1 \\ -2 & 3 & 2 \\ 4 & -1 & 5 \end{pmatrix} = \begin{pmatrix} -8 & 3 & -1 \\ 0 & 5 & 9 \\ 14 & -6 & 13 \end{pmatrix}.$$

Vemos, pues, que el producto de las matrices depende del orden de los factores, es decir, que la *multiplicación de las matrices no es conmutativa*. Por cierto, esto era de esperar, aunque sólo sea por el hecho de que, en la definición de la matriz C, dada anteriormente mediante la fórmula (3), las matrices A y B no figuran de un modo equivalente: en A se toman las filas, mientras que en B, las columnas.

Se pueden señalar, para todos los n, comenzando desde $n = 2$, ejemplos de matrices de n-ésimo orden no conmutables, o sea, de matrices cuyo producto se altera al permutar los factores (las matrices de segundo orden en el ejemplo 1) no son conmutables). Por otra parte, *dos matrices pueden ser ocasionalmente conmutables*, como muestra el siguiente ejemplo:

$$\begin{pmatrix} 7 & -12 \\ -4 & 7 \end{pmatrix} \cdot \begin{pmatrix} 26 & 45 \\ 15 & 26 \end{pmatrix} = \begin{pmatrix} 26 & 45 \\ 15 & 26 \end{pmatrix} \cdot \begin{pmatrix} 7 & -12 \\ -4 & 7 \end{pmatrix} = \begin{pmatrix} 2 & 3 \\ 1 & 2 \end{pmatrix}.$$

El producto de las matrices es asociativo; por consiguiente, se puede hablar del producto, unívocamente determinado, de cualquier número finito de matrices de n-ésimo orden, tomadas (en virtud de que el producto no es conmutativo) en un orden determinado.

Demostración. Sean dadas tres matrices arbitrarias de n-ésimo orden, A, B, y C. Escribámoslas del modo abreviado siguiente, donde se indica la forma general de sus elementos: $A = (a_{ij})$, $B = (b_{ij})$, $C = (c_{ij})$. Introduzcamos luego las siguientes notaciones:

$$AB = U = (u_{ij}), \qquad BC = V = (v_{ij}),$$
$$(AB)\,C = S = (s_{ij}), \qquad A\,(BC) = T = (t_{ij}).$$

Tenemos que demostrar que se cumple la igualdad $(AB)\,C = A\,(BC)$, es decir, $S = T$. Sin embargo,

$$u_{il} = \sum_{k=1}^{n} a_{ik} b_{kl}, \quad v_{kj} = \sum_{l=1}^{n} b_{kl} c_{lj},$$

de donde, en virtud de las igualdades $S = UC$ y $T = AV$,

$$s_{ij} = \sum_{l=1}^{n} u_{il} c_{lj} = \sum_{l=1}^{n} \sum_{k=1}^{n} a_{ik} b_{kl} c_{lj},$$

$$t_{ij} = \sum_{k=1}^{n} a_{ik} v_{kj} = \sum_{k=1}^{n} \sum_{l=1}^{n} a_{ik} b_{kl} c_{lj},$$

o sea, $s_{ij} = t_{ij}$ para $i, j = 1, 2, \ldots, n$.

Para el estudio ulterior de las propiedades del producto de las matrices se necesita el empleo de los determinantes. Además, para abreviar, convendremos en designar con $|A|$ el determinante de la matriz A. Si en cada uno de los ejemplos considerados anteriormente el lector calcula los determinantes de las matrices que se multiplican y compara el producto de estos determinantes con el determinante del producto de las matrices dadas, puede observar una ley bastante curiosa que se expresa con el siguiente importante **teorema sobre el producto de los determinantes:**

El determinante del producto de varias matrices de n-ésimo orden es igual al producto de los determinantes de estas matrices.

Es suficiente demostrar este teorema para el caso de dos matrices. Sean dadas las matrices de n-ésimo orden $A = (a_{ij})$ y $B = (b_{ij})$, y sea $AB = C = (c_{ij})$. Formemos el siguiente determinante auxiliar Δ de orden $2n$: en su ángulo superior de la izquierda colocamos la matriz A, en el ángulo inferior de la derecha, la matriz B, todo el ángulo superior de la derecha lo ocupamos con ceros. Finalmente, formamos la diagonal principal del ángulo inferior de la izquierda con el número -1, ocupando todos los demás lugares también con ceros. Por consiguiente, el determinante Δ tiene la forma siguiente:

$$\Delta = \begin{vmatrix} a_{11} & a_{12} & \dots & a_{1n} & 0 & 0 & \dots & 0 \\ a_{21} & a_{22} & \dots & a_{2n} & 0 & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} & 0 & 0 & \dots & 0 \\ -1 & 0 & \dots & 0 & b_{11} & b_{12} & \dots & b_{1n} \\ 0 & -1 & \dots & 0 & b_{21} & b_{22} & \dots & b_{2n} \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & -1 & b_{n1} & b_{n2} & \dots & b_{nn} \end{vmatrix}.$$

La aplicación del teorema de Laplace al determinante — su desarrollo por los menores de las primeras n filas — nos lleva a la siguiente igualdad:

$$\Delta = |A| \cdot |B|.$$

Procuremos, a su vez, transformar el determinante Δ de tal modo que, sin cambiar su valor, todos los elementos b_{ij}, $i, j = 1, 2, \dots, n$, queden sustituidos por ceros. Con este fin, agreguemos a la $(n+1)$-ésima columna del determinante Δ su primera columna, multiplicada por b_{11}, su segunda columna, multiplicada por b_{21}, etc., y finalmente, su n-ésima columna, multiplicada por b_{n1}. Después, agreguemos a la $(n+2)$-ésima columna del determinante Δ la primera columna, multiplicada por b_{12}, la segunda columna, multiplicada por b_{22}, etc. En general, agreguemos a la $(n+j)$-

ésima columna del determinante Δ, donde $j = 1, 2, \ldots, n$, la suma de las primeras n columnas, tomadas con los coeficientes $b_{1j}, b_{2j}, \ldots, b_{nj}$, respectivamente.

Fácilmente se ve que estas transformaciones, no alterando el determinante, dan lugar a la sustitución de todos los elementos b_{ij} por ceros. A la vez, en lugar de los ceros que figuraban en el ángulo superior de la derecha del determinante, aparecerán los números siguientes: en la intersección de la i-ésima fila y $(n + j)$-ésima columna del determinante, $i, j = 1, 2, \ldots, n$, estará ahora el número $a_{i1}b_{1j} + a_{i2}b_{2j} + \ldots + a_{in}b_{nj}$, que en virtud de (3) es igual al elemento c_{ij} de la matriz $C = AB$. Por consiguiente, el ángulo superior de la derecha del determinante lo ocupa ahora la matriz C:

$$\Delta = \begin{vmatrix} a_{11} & a_{12} & \ldots & a_{1n} & c_{11} & c_{12} & \ldots & c_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} & c_{21} & c_{22} & \ldots & c_{2n} \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ a_{n1} & a_{n2} & \ldots & a_{nn} & c_{n1} & c_{n2} & \ldots & c_{nn} \\ -1 & 0 & \ldots & 0 & 0 & 0 & \ldots & 0 \\ 0 & -1 & \ldots & 0 & 0 & 0 & \ldots & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & \ldots & -1 & 0 & 0 & \ldots & 0 \end{vmatrix}$$

Apliquemos otra vez más el teorema de Laplace, desarrollando el determinante por los menores de las últimas n columnas. Como el menor complementario para el menor $|C|$ es igual a $(-1)^n$, y el menor $|C|$ está situado en las filas cuyos números de orden son $1, 2, \ldots, n$ y en las columnas cuyos números de orden son $n + 1, n + 2, \ldots, 2n$, aplicando la igualdad

$$1 + 2 + \ldots + n + (n + 1) + (n + 2) + \ldots + 2n = 2n^2 + n,$$

se tiene

$$\Delta = (-1)^{2n^2+n}(-1)^n |C| = (-1)^{2(n^2+n)} |C|$$

o bien, como el número $2(n^2 + n)$ es par,

$$\Delta = |C|. \tag{5}$$

Finalmente, de (4) y (5) se deduce la igualdad que queríamos demostrar

$$|C| = |A| \cdot |B|.$$

El teorema sobre el producto de los determinantes podría haber sido demostrado también sin la utilización del teorema de Laplace. El lector hallará una de estas demostraciones al final del § 16.

§ 14. Matriz inversa

Una matriz cuadrada se llama *degenerada* (o *singular*), si su determinante es igual a cero, y *no degenerada* (o *no singular*)*, en el caso contrario. Correspondientemente, una transformación lineal de las indeterminadas se llama *degenerada* o *no degenerada*, según que el determinante de los coeficientes de esta transformación sea igual a cero o no. Del teorema demostrado al final del párrafo anterior, se deduce la afirmación siguiente:

El producto de matrices, al menos una de las cuales es degenerada, es también una matriz degenerada.

El producto de cualesquiera matrices no degeneradas también es una matriz no degenerada. De aquí se deduce, en virtud de la relación existente entre el producto de las matrices y la realización consecutiva de las transformaciones lineales, la proposición siguiente: *el resultado de la realización consecutiva de unas cuantas transformaciones lineales será una transformación no degenerada cuando, y sólo cuando, todas las transformaciones dadas sean no degeneradas.*

En el producto de las matrices, el papel de la unidad lo desempeña la matriz

$$E=\begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ . & . & . & . \\ 0 & 0 & & 1 \end{pmatrix},$$

que es además conmutable con cualquier matriz A del orden dado,

$$AE=EA=A. \qquad (1)$$

Estas igualdades se demuestran aplicando directamente la regla de multiplicación de las matrices, o basándose en la observación de que la matriz unidad corresponde a la transformación lineal *idéntica* de las indeterminadas

$$\begin{aligned} x_1 &= y_1, \\ x_2 &= \quad y_2, \\ &\dots\dots\dots \\ x_n &= \qquad y_n, \end{aligned}$$

cuya realización, antes o después de cualquier otra transformación lineal, no cambia, evidentemente, esta última.

Obsérvese que la *matriz E es la única que satisface a la condición* (1) *para cualquier matriz A*. En efecto, si existiese otra matriz E' con esta propiedad, tendríamos que

$$E'E=E', \quad E'E=E,$$

de donde, $E'=E$.

* También se llama matriz regular. (*Nota del T.*)

El problema de la existencia de la *matriz inversa* para una matriz dada A es más complicado. Como el producto de las matrices no es conmutativo, hablaremos ahora de la matriz inversa a la *derecha*, o sea, de una matriz A^{-1} tal, que el producto de la matriz A por esta matriz a la derecha es igual a la matriz unidad,

$$AA^{-1} = E. \tag{2}$$

Si la matriz A es degenerada y existiese la matriz A^{-1}, el producto que figura en el primer miembro de la igualdad (2) sería, como ya sabemos, una matriz degenerada. En realidad, la matriz E que figura en el segundo miembro de esta igualdad no es degenerada, puesto que su determinante es igual a la unidad. Por lo tanto, una matriz degenerada no puede tener matriz inversa a la derecha. Estos mismos razonamientos muestran que ésta no puede tener tampoco matriz inversa a la izquierda, *no existiendo, por lo tanto, matriz inversa para una matriz degenerada.*

Refiriéndonos al caso de una matriz no degenerada, introduzcamos primero el siguiente concepto auxiliar. Sea dada una matriz de n-ésimo orden

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{pmatrix}.$$

La matriz

$$A^* = \begin{pmatrix} A_{11} & A_{21} & \dots & A_{n1} \\ A_{12} & A_{22} & \dots & A_{n2} \\ \dots & \dots & \dots & \dots \\ A_{1n} & A_{2n} & \dots & A_{nn} \end{pmatrix},$$

formada por los complementos algebraicos de los elementos de la matriz A, donde el complemento algebraico del elemento a_{ij} está situado en la intersección de la j-ésima fila y de la i-ésima columna, se llama matriz *adjunta* de la matriz A.

Hallemos los poductos AA^* y A^*A. Aplicando la fórmula estudiada en el § 6, sobre el desarrollo de un determinante por los elementos de una fila o columna, y también el teorema del § 7, sobre la suma de los productos de los elementos de cualquier fila (o columna) de un determinante por los complementos algebraicos de los elementos correspondientes de otra fila (columna), y designando con d el determinante de la matriz A,

$$d = |A|,$$

obtenemos las siguientes igualdades:

$$AA^* = A^*A = \begin{pmatrix} d & 0 \dots 0 \\ 0 & d \dots 0 \\ \dots & \dots \\ 0 & 0 \dots d \end{pmatrix}. \qquad (3)$$

De aquí se deduce que, si la *matriz A no es degenerada, su matriz adjunta A^* tampoco lo será, siendo además, el determinante d^* de la matriz A^* igual a la $(n-1)$-ésima potencia del determinante d de la matriz A.*

En efecto, pasando de las igualdades (3) a la igualdad entre los determinantes, obtenemos:

$$dd^* = d^n,$$

de donde, como $d \neq 0$,

$$d^* = d^{n-1}*).$$

Ahora es fácil demostrar la existencia de una matriz inversa para cualquier matriz A **no degenerada,** y hallar su forma. Obsérvese primero que si se considera el producto de dos matrices AB y se dividen por un mismo número d todos los elementos de uno de los factores, por ejemplo B, entonces, todos los elementos del producto AB también se dividirán por este mismo número: para la demostración solamente hay que recordar la definición del producto de las matrices. Por lo tanto, si

$$d = |A| \neq 0,$$

de las igualdades (3) se deduce, que la *inversa de la matriz A es la matriz que resulta de la matriz adjunta A^* al dividir todos sus elementos por el número d*:

$$A^{-1} = \begin{pmatrix} \frac{A_{11}}{d} & \frac{A_{21}}{d} & \dots & \frac{A_{n1}}{d} \\ \frac{A_{12}}{d} & \frac{A_{22}}{d} & \dots & \frac{A_{n2}}{d} \\ \dots & \dots & \dots & \dots \\ \frac{A_{1n}}{d} & \frac{A_{2n}}{d} & \dots & \frac{A_{nn}}{d} \end{pmatrix}.$$

En efecto, de (3) se deducen las igualdades

$$AA^{-1} = A^{-1}A = E. \qquad (4)$$

Subrayemos una vez más que en la i-ésima **fila** de la matriz A^{-1} figuran los complementos algebraicos de los elementos de la i-ésima **columna** del determinante $|A|$, divididos por $d = |A|$.

* Se podría demostrar que si la matriz A es degenerada, su matriz adjunta A^* también lo es, teniendo además un rango no superior al número 1.

Es fácil demostrar que la *matriz A^{-1} es la única que satisface a la condición* (4) *para una matriz dada A, no degenerada*. En efecto, si la matriz C es tal que

$$AC=CA=E,$$

entonces,

$$CAA^{-1}=C\,(AA^{-1})=CE=C,$$
$$CAA^{-1}=(CA)\,A^{-1}=EA^{-1}=A^{-1},$$

de donde $C=A^{-1}$.

De (4) y del teorema sobre el producto de los determinantes, se deduce que el *determinante de la matriz A^{-1} es igual a $\frac{1}{|A|}$. Así, pues, esta matriz tampoco es degenerada; la inversa para ella es la misma matriz A.*

Si se dan ahora las matrices cuadradas A y B de n-ésimo orden, de las cuales A no es degenerada, mientras que B es arbitraria, podemos efectuar la división **por la derecha y por la izquierda** de B por A, es decir, resolver las ecuaciones matriciales

$$AX=B,\ YA=B. \tag{5}$$

Para esto, en virtud de la asociatividad del producto de las matrices, es suficiente hacer

$$X=A^{-1}B,\ Y=BA^{-1};$$

como el producto de las matrices no es conmutativo, por lo general, estas soluciones de las ecuaciones (5) serán diferentes.

Ejemplos. 1) Se da la matriz

$$A=\begin{pmatrix} 3 & -1 & 0 \\ -2 & 1 & 1 \\ 2 & -1 & 4 \end{pmatrix}.$$

Su determinante $|A|=5$, por consiguiente, la matriz inversa A^{-1} existe:

$$A^{-1}=\begin{pmatrix} 1 & \frac{4}{5} & -\frac{1}{5} \\ 2 & \frac{12}{5} & -\frac{3}{5} \\ 0 & \frac{1}{5} & \frac{1}{5} \end{pmatrix}.$$

2) Se dan las matrices

$$A=\begin{pmatrix} 3 & 2 \\ 4 & 3 \end{pmatrix},\ B=\begin{pmatrix} -1 & 7 \\ 3 & 5 \end{pmatrix}.$$

La matriz A no es degenerada, y además,

$$A^{-1}=\begin{pmatrix} 3 & -2 \\ -4 & 3 \end{pmatrix},$$

por lo tanto, las soluciones de las ecuaciones $AX=B$, $YA=B$ serán las matrices

$$X=\begin{pmatrix} 3 & -2 \\ -4 & 3 \end{pmatrix}\cdot\begin{pmatrix} -1 & 7 \\ 3 & 5 \end{pmatrix}=\begin{pmatrix} -9 & 11 \\ 13 & -13 \end{pmatrix},$$

$$Y=\begin{pmatrix} -1 & 7 \\ 3 & 5 \end{pmatrix}\cdot\begin{pmatrix} 3 & -2 \\ -4 & 3 \end{pmatrix}=\begin{pmatrix} -31 & 23 \\ -11 & 9 \end{pmatrix}.$$

Multiplicación de matrices rectangulares. El producto de las matrices definido en el párrafo anterior solamente para las matrices cuadradas de igual orden, puede generalizarse también para el caso de matrices rectangulares A y B, siempre que sea posible aplicar la fórmula (3) del párrafo anterior, o sea, cuando cada fila de la matriz A contenga tantos elementos como haya en cada columna de la matriz B. En otras palabras, *se puede hablar del producto de las matrices rectangulares A y B cuando el número de columnas de la matriz A es igual al número de filas de la matriz B. En este caso, el número de filas de la matriz AB es igual al número de filas de la matriz A, y el número de columnas de la matriz AB es igual al número de columnas de la matriz B.*

Ejemplos.

1) $$\begin{pmatrix} 5 & -1 & 3 & 1 \\ 2 & 0 & -1 & 4 \end{pmatrix}\cdot\begin{pmatrix} -1 & 3 & 0 \\ -2 & 1 & 1 \\ 3 & 0 & -2 \\ 4 & 1 & 2 \end{pmatrix}=\begin{pmatrix} 10 & 15 & -5 \\ 11 & 10 & 10 \end{pmatrix}.$$

2) $$\begin{pmatrix} 0 & -3 & 1 \\ 2 & 1 & 5 \\ -4 & 0 & -2 \end{pmatrix}\cdot\begin{pmatrix} 3 \\ -2 \\ 2 \end{pmatrix}=\begin{pmatrix} 8 \\ 14 \\ -16 \end{pmatrix}.$$

3) $$(5 \quad 1 \quad 0 \quad -3)\cdot\begin{pmatrix} 2 & 0 \\ 1 & -4 \\ 3 & 1 \\ 0 & -1 \end{pmatrix}=(11 \quad -1).$$

El producto de las matrices rectangulares se puede ligar con la ejecución consecutiva de las transformaciones lineales de las indeterminadas solamente si en la definición de estas últimas no se insiste en que se conserve el número de indeterminadas en la transformación lineal.

Repitiendo palabra por palabra la demostración dada anteriormente para el caso de las matrices cuadradas, se comprueba fácilmente que la *ley asociativa se cumple también para el producto de matrices rectangulares.*

Ahora utilizaremos el producto de las matrices rectangulares y las propiedades de la matriz inversa para **deducir de nuevo la regla de Cramer,** evitando los complicados cálculos que se realizaron en el § 7. *Sea dado un sistema de n ecuaciones lineales con n incógnitas:*

$$\left.\begin{array}{l} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n = b_1, \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n = b_2, \\ \dots\dots\dots\dots\dots\dots \\ a_{n1}x_1 + a_{n2}x_2 + \ldots + a_{nn}x_n = b_n, \end{array}\right\} \qquad (6)$$

donde el determinante del sistema es diferente de cero. Designemos con A la matriz de los coeficientes del sistema (6); esta matriz no es degenerada, ya que, por la suposición hecha, $d = |A| \neq 0$. Designemos con X *la columna de las incógnitas; con* B, *la columna* de los términos independientes del sistema (6), es decir,

$$X = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}, \quad B = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix}.$$

El producto AX tiene sentido, puesto que el número de columnas de la matriz A es igual al número de filas de la matriz X, y además, este producto será una columna, formada por los primeros miembros de las ecuaciones del sistema (6). Por lo tanto, el sistema (6) puede escribirse en forma de una ecuación matricial

$$AX = B. \qquad (7)$$

Multiplicando a la izquierda ambos miembros de la ecuación (7) por la matriz A^{-1}, cuya existencia es consecuencia de que la matriz cuadrada A no es degenerada, obtenemos:

$$X = A^{-1}B. \qquad (8)$$

El producto que figura en el segundo miembro de esta igualdad es una matriz de una sola columna; su j-ésimo elemento es igual a la suma de los productos de los elementos de la j-ésima fila de la matriz A^{-1} por los elementos correspondientes de la matriz B, es decir, es igual al número

$$\frac{A_{1j}}{d} b_1 + \frac{A_{2j}}{d} b_2 + \ldots + \frac{A_{nj}}{d} b_n = \frac{1}{d}(A_{1j}b_1 + A_{2j}b_2 + \ldots + A_{nj}b_n).$$

Sin embargo, la expresión que figura entre paréntesis en el segundo miembro de la igualdad es el desarrollo por los elementos de la j-ésima columna del determinante d_j, que se obtiene sustituyendo la j-ésima columna del determinante d por la columna B. Por lo tanto, las fórmulas (8) son equivalentes a las fórmulas (3) del § 7, que expresan la solución del sistema (6) obtenida por la regla de Cramer.

Queda por demostrar que los valores obtenidos de las incógnitas forman verdaderamente una solución del sistema (6). Con este fin, es suficiente poner la expresión (8) en la ecuación matricial (7), lo que da lugar, evidentemente, a la identidad $B = B$.

El rango del producto de las matrices. En el caso de matrices degeneradas, el teorema del producto de los determinantes no nos lleva a ningún enunciado más de que tal producto también es degenerado, a pesar de que las matrices cuadradas degeneradas se pueden diferenciar también por su rango. Obsérvese que no existe una dependencia absolutamente determinada entre los rangos de los factores y el rango del producto, como se muestra en los ejemplos siguientes:

$$\begin{pmatrix} 2 & 0 \\ 0 & 0 \end{pmatrix} \cdot \begin{pmatrix} 3 & 0 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 6 & 0 \\ 0 & 0 \end{pmatrix},$$

$$\begin{pmatrix} 2 & 0 \\ 0 & 0 \end{pmatrix} \cdot \begin{pmatrix} 0 & 0 \\ 0 & 3 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix};$$

en ambos casos se multiplican matrices de rango 1. Sin embargo, en un caso el producto tiene el rango 1, mientras que en el otro, el rango es 0. Subsiste solamente el siguiente teorema, que no sólo es justo para las matrices cuadradas, sino también para las matrices rectangulares:

El rango del producto de varias matrices no es superior al rango de cada uno de los factores.

Es suficiente demostrar este teorema para el caso de dos factores. Sean dadas las matrices A y B, para las cuales tiene sentido el producto AB; emplearemos la notación $AB = C$. Veamos la fórmula (3) del § 13, que da la expresión de los elementos de la matriz C. Tomando esta fórmula para un k dado y todos los i posibles, ($i = 1, 2, \ldots$) obtenemos que la k-ésima columna de la matriz C representa una suma de todas las columnas de la matriz A, tomadas con ciertos coeficientes (precisamente con los coeficientes b_{1k}, b_{2k}, ...). De este modo, queda demostrado que el sistema de columnas de la matriz C se expresa linealmente mediante el sistema de columnas de la matriz A, y, por consiguiente, como se ha demostrado en el § 9, el rango del primer sistema es menor o igual al rango del segundo sistema; en otras palabras, el rango de la matriz C no es mayor que el rango de la matriz A. Por otra parte, como de la misma fórmula (3) del

§ 13, para un i dado y todos los k, se deduce que toda i-ésima fila de la matriz C es combinación lineal de las filas de la matriz B, con razonamientos análogos obtenemos que el rango de C no es mayor que el rango de B.

Cuando uno de los factores representa una matriz cuadrada no degenerada, se obtiene un resultado más exacto.

El rango del producto a la derecha y a la izquierda de una matriz A por una matriz cuadrada no degenerada Q, es igual al rango de la matriz A.

Sea, por ejemplo,

$$AQ = C. \tag{9}$$

Del teorema precedente se deduce que el rango de la matriz C no es mayor que el rango de la matriz A. Por otra parte, multiplicando a la derecha la igualdad (9) por Q^{-1}, llegamos a la igualdad

$$A = CQ^{-1},$$

y, por consiguiente, otra vez por el teorema precedente, el rango de A no es mayor que el rango de C. Comparando estos dos resultados obtenemos la coincidencia de los rangos de las matrices A y C.

§ 15. Suma de matrices y multiplicación de una matriz por un número

Para las matrices cuadradas de orden n, la **suma** se define del modo siguiente:

Se llama *suma* $A + B$ *de dos matrices cuadradas* $A = (a_{ij})$ *y* $B = (b_{ij})$ *de orden* n, a una matriz $C = (c_{ij})$ tal, que cualquier elemento de ella es igual a la suma de los elementos correspondientes de las matrices A y B;

$$c_{ij} = a_{ij} + b_{ij} \text{*}$$

Es evidente, que la suma de matrices definida es conmutativa y asociativa. Para ella existe la operación inversa: la resta, llamándose diferencia de las matrices A y B a la matriz formada por las diferencias de los elementos correspondientes de las matrices dadas. En este caso, el papel de cero lo desempeña la *matriz nula*, compuesta totalmente de ceros; a continuación, esta matriz se designará con el símbolo 0: no hay peligro de confundir la matriz nula con el número cero.

La suma de las matrices cuadradas y el producto de éstas, definido en el § 13, están ligados con las leyes distributivas.

---

* Por supuesto, se podría definir también el producto de matrices multiplicando los elementos correspondientes. Sin embargo, esta multiplicación, a diferencia de la que se definió en el § 13, carecería de aplicaciones serias.

En efecto, sean dadas tres matrices de orden n, $A = (a_{ij})$, $B = (b_{ij})$, $C = (c_{ij})$. Entonces, para cualesquiera i y j, se cumple la igualdad

$$\sum_{s=1}^{n} (a_{is} + b_{is}) c_{sj} = \sum_{s=1}^{n} a_{is}c_{sj} + \sum_{s=1}^{n} b_{is}c_{sj}.$$

El primer miembro de esta igualdad es el elemento situado en la i-ésima fila y j-ésima columna de la matriz $(A + B) C$, el segundo miembro es el elemento situado en el mismo lugar, pero en la matriz $AC + BC$. Con esto, queda demostrada la igualdad

$$(A + B) C = AC + BC.$$

La igualdad $C (A + B) = CA + CB$ se demuestra del mismo modo. Como el producto de matrices no es conmutativo, se deben demostrar estas dos leyes distributivas.

Introduzcamos la siguiente definición de producto de una matriz por un número.

Se llama *producto* kA *de una matriz cuadrada* $A = (a_{ij})$ *por el número* k, a la matriz $A' = (a'_{ij})$ que se obtiene multiplicando por k todos los elementos de la matriz A:

$$a'_{ij} = ka_{ij}.$$

En el párrafo anterior ya tratamos un ejemplo de multiplicación de una matriz por un número: si la matriz A no es degenerada, siendo $|A| = d$, su matriz inversa A^{-1} y su matriz adjunta A^* están ligadas por la igualdad

$$A^{-1} = d^{-1}A^*.$$

Como ya sabemos, toda matriz cuadrada de orden n se puede considerar como un vector de n^2 dimensiones, siendo biunívoca esta correspondencia entre las matrices y los vectores. En este caso, las operaciones definidas de suma de matrices y de producto de una matriz por un número, se convierten en la suma de vectores y en el producto de un vector por un número. Por lo tanto, el *conjunto de las matrices cuadradas de orden n se puede considerar como un espacio vectorial de n^2 dimensiones.*

De aquí se deduce el cumplimiento de las igualdades siguientes (aquí, A, B son matrices de orden n; k, l son unos números; 1 es el número uno):

$$k (A + B) = kA + kB, \tag{1}$$

$$(k + l) A = kA + lA, \tag{2}$$

$$k (lA) = (kl) A, \tag{3}$$

$$1 \cdot A = A. \tag{4}$$

Las propiedades (1) y (2) ligan el producto de una matriz por un número con la suma de matrices. Al mismo tiempo, existe una ligazón importante entre el producto de una matriz por un número y el producto de las matrices mismas, que se expresa así:

$$(kA)B = A(kB) = k(AB), \qquad (5)$$

o sea, *si en el producto de las matrices, uno de los factores se multiplica por el número k, todo el producto queda multiplicado por k.*

En efecto, sean dadas las matrices $A = (a_{ij})$ y $B = (b_{ij})$ y el número k. Entonces, para cualesquiera i y j, se tiene:

$$\sum_{s=1}^{n} (ka_{is}) b_{sj} = k \sum_{s=1}^{n} a_{is} b_{sj}.$$

Pero, el primer miembro de esta igualdad es el elemento situado en la i-ésima fila y j-ésima columna de la matriz $(kA)B$, y el segundo miembro es el elemento situado en el mismo sitio en la matriz $k(AB)$. Con esto queda demostrada la igualdad

$$(kA)B = k(AB).$$

La igualdad $A(kB) = k(AB)$ se demuestra del mismo modo.

La *multiplicación de una matriz por un número permite introducir* un nuevo método de expresión de las matrices. Designemos con E_{ij} *la matriz en la que, en la intersección de la i-ésima fila y j-ésima columna* figura la unidad, mientras que todos los demás elementos *son iguales a cero*. Haciendo $i = 1, 2, \ldots, n$ y $j = 1, 2, \ldots, n$, obtenemos n^2 matrices de éstas, E_{ij}, que, como fácilmente se comprueba, están ligadas por la siguiente tabla de multiplicar:

$$E_{is}E_{sj} = E_{ij}, \quad E_{is}E_{tj} = 0 \text{ para } s \neq t.$$

La matriz kE_{ij} solamente se diferencia de la matriz E_{ij} en que en ella figura el número k en la intersección de la i-ésima fila y j-ésima columna. Teniendo esto en cuenta y aplicando la definición de suma de matrices, obtenemos la siguiente expresión para una matriz cuadrada arbitraria A:

$$A = \begin{pmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \cdot & \cdot & \cdot & \cdot \\ a_{n1} & a_{n2} & \ldots & a_{nn} \end{pmatrix} = \sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij} E_{ij}, \qquad (6)$$

poseyendo, evidentemente, la matriz A una sola expresión de la forma (6).

La matriz kE, donde E es la matriz unidad, según la definición del producto de una matriz por un número, tiene la forma siguiente:

$$kE = \begin{pmatrix} k & & & 0 \\ & k & & \\ & & \ddots & \\ 0 & & & k \end{pmatrix},$$

o sea, en la diagonal principal figura un mismo número k, mientras que todos los elementos situados fuera de esta diagonal son iguales a cero. Tales matrices se llaman *escalares.*

La definición de la suma de matrices conduce a la siguiente igualdad

$$kE + lE = (k + l) E. \tag{7}$$

Por otra parte, aplicando la definición del producto de matrices o basándose en la igualdad (5), obtenemos:

$$kE \cdot lE = (kl) E. \tag{8}$$

El producto de una matriz A por un número k se puede interpretar como el producto obtenido al multiplicar A por la matriz escalar kE, en el sentido de la multiplicación de matrices. En efecto, según (5)

$$(kE) A = A (kE) = kA.$$

De aquí se deduce también, que *toda matriz escalar es conmutable con cualquier matriz A.* Es de gran importancia tener en cuenta que las matrices escalares son las únicas que poseen esta propiedad:

Si una matriz $C = (c_{ij})$ de n-ésimo orden es conmutable con cualquier matriz del mismo orden, la matriz C es escalar.

En efecto, supongamos que $i \neq j$ y consideremos los productos CE_{ij} y $E_{ij}C$ (véase más arriba la definición de la matriz E_{ij}) que, por hipótesis, son iguales entre sí. Fácilmente se observa que todas las columnas de la matriz CE_{ij}, menos la j-ésima, se componen de ceros, y que la j-ésima columna coincide con la i-ésima columna de la matriz C; en particular, en la intersección de la i-ésima fila y j-ésima columna de la matriz CE_{ij} está situado el elemento c_{ii}. Análogamente, todas las filas de la matriz $E_{ij}C$, menos la i-ésima, se componen de ceros, y la i-ésima fila coincide con la j-ésima fila de la matriz C; en la intersección de la i-ésima fila y j-ésima columna de la matriz $E_{ij}C$ está situado el elemento c_{jj}. Aplicando la igualdad $CE_{ij} = E_{ij}C$, obtenemos que: $c_{ii} = c_{jj}$ (como elementos situados en lugares iguales de matrices que son iguales entre sí); o sea, todos los elementos de la diagonal principal de la matriz C son iguales entre sí. Por otra parte, en la intersección de la j-ésima fila y j-ésima columna de la matriz CE_{ij} está el elemento c_{ji}; pero,

en la matriz $E_{ij}C$, en este sitio está el cero (en virtud de que $i \neq j$), de donde: $c_{ji} = 0$, es decir, cualquier elemento de la matriz C, situado fuera de la diagonal principal, es igual a cero. El teorema está demostrado.

§ 16. Construcción axiomática de la teoría de los determinantes

Un determinante de n-ésimo orden representa un número, unívocamente determinado por una matriz cuadrada dada de n-ésimo orden. La definición de este concepto, expuesta en el § 4, da la regla, según la cual el determinante se expresa mediante los elementos de la matriz dada. Sin embargo, esta definición constructiva se puede sustituir por una axiomática; mejor dicho, entre las propiedades del determinante establecidas en los §§ 4 y 6, se pueden indicar algunas, de modo que la única función de la matriz con valores reales que posea estas propiedades sea su determinante.

La definición más simple de este género consiste en la utilización de los desarrollos del determinante por los elementos de una fila. Consideremos las matrices cuadradas de cualesquiera órdenes y supongamos que a cada matriz M de éstas se le ha puesto en correspondencia un número d_M, cumpliéndose las condiciones siguientes:

1) Si la matriz M es de primer orden, o sea, que consta de un elemento a, entonces $d_M = a$.

2) Si los elementos $a_{11}, a_{12}, \ldots, a_{1n}$ forman la primera fila de la matriz M de n-ésimo orden y si se ha designado con M_i, $i = 1, 2, \ldots, n$, la matriz de $(n-1)$-ésimo orden que queda después de suprimir en M la primera fila y la i-ésima columna, entonces

$$d_M = a_{11}d_{M_1} - a_{12}d_{M_2} + a_{13}d_{M_3} - \ldots + (-1)^{n-1} a_{1n}d_{M_n}.$$

Por lo tanto, para cualquier matriz M, el número d_M es igual al determinante de esta matriz. La demostración de esta afirmación la dejamos a cuenta del lector. Esta se efectúa por el método de inducción sobre n y se basa en los resultados del § 6.

Mucho más interesantes son otras formas de definición axiomática de los determinantes, que se refieren al caso de un orden dado n y se basan en algunas de las propiedades simples de los determinantes, establecidas en el § 4. Examinaremos ahora una de estas definiciones.

Supongamos que a cada matriz cuadrada M de n-ésimo orden se pone en correspondencia un número d_M, cumpliéndose las condiciones siguientes:

I. *Si una de las filas de la matriz M se multiplica por un número k, el número d_M queda multiplicado por k.*

II. *El número d_M no varía si a una de las filas de la matriz M se le agrega otra fila de esta matriz.*

III. *Si E es la matriz unidad, entonces $d_E = 1$.*

Demostremos que *para cualquier matriz M, el número d_M es igual al determinante de esta matriz.*

Establezcamos primero, partiendo de las condiciones I—III, unas propiedades del número d_M, que son análogas a las propiedades correspondientes de los determinantes.

(1) *Si una de las filas de la matriz M consta de ceros, entonces $d_M = 0$.*

En efecto, multiplicando la fila compuesta de ceros por el número 0, la matriz no varía. Sin embargo, en virtud de la condición I, el número d_M adquiere el factor 0, de donde,

$$d_M = 0 \cdot d_M = 0.$$

(2) *El número d_M no varía si a la i-ésima fila de la matriz M se le agrega la j-ésima fila, $j \neq i$, multiplicada por el número k.*

Si $k = 0$, todo está demostrado. Si $k \neq 0$, multiplicamos la j-ésima fila por k y obtenemos la matriz M', para la que $d_{M'} = kd_M$, en virtud de la condición I. Después agregamos a la i-ésima fila de la matriz M' su j-ésima fila y obtenemos la matriz M'', y en virtud de la condición II, se tiene $d_{M''} = d_{M'}$. Finalmente, multiplicamos la j-ésima fila de la matriz M'' por el número k^{-1}, obteniendo la matriz M''', que en realidad resulta de M mediante la transformación indicada en el enunciado de la propiedad que estamos demostrando; además.

$$d_{M'''} = k^{-1}d_{M''} = k^{-1}d_{M'} = k^{-1} \cdot kd_M = d_M.$$

(3) *Si las filas de la matriz M son linealmente dependientes, entonces $d_M = 0$.*

En efecto, si una de las filas, por ejemplo la i-ésima, es combinación lineal de las otras filas, entonces aplicando unas cuantas veces la transformación (2), se puede sustituir la i-ésima fila por ceros. La transformación (2) no altera el número d_M, por lo cual, en virtud de la propiedad (1), se tiene $d_M = 0$.

(4) *Si la i-ésima fila de la matriz M es la suma de dos vectores β y γ, y si las matrices M' y M'' se obtienen de la matriz M sustituyendo su i-ésima fila por los vectores β y γ, respectivamente, entonces*

$$d_M = d_{M'} + d_{M''}.$$

En efecto, sea S el sistema de todas las filas de la matriz M, excluyendo la i-ésima. Si existe en S una dependencia lineal, las filas de cada una de las matrices M, M' y M'' son linealmente dependientes, de donde, por la propiedad (3), $d_M = d_{M'} = d_{M''} = 0$.

De aquí se deduce la validez de la propiedad en cuestión. Si el sistema S *constituido de* $n - 1$ vectores es linealmente independiente, entonces, como muestran los resultados del § 9, éste se puede completar con un vector α de modo que resulte un sistema linealmente independiente maximal de vectores del espacio de n dimensiones. Los vectores β y γ se pueden expresar linealmente mediante este sistema. Supongamos que el vector α figura en esta expresión con los coeficientes k y l, respectivamente; por consiguiente, en la expresión del vector $\beta + \gamma$, o sea, en la i-ésima fila de la matriz M, el vector α figurará con el coeficiente $k + l$. Ahora se pueden transformar las matrices M, M' y M'', restando de sus i-ésimas filas ciertas combinaciones lineales de las otras filas, de modo que en las i-ésimas filas resulten los vectores $(k + l)\alpha$, $k\alpha$, y $l\alpha$, respectivamente. En consecuencia, designando con M^0 la matriz que resulta de la matriz M sustituyendo su i-ésima fila por el vector α, y, teniendo en cuenta las propiedades (2) y I, llegamos a las igualdades:

$$d_M = (k + l)\, d_{M^0},\ \ d_{M'} = k d_{M^0},\ \ d_{M''} = l d_{M^0}.$$

Con esto, la propiedad (4) queda demostrada.

(5) *Si la matriz* $\overline{M}$ *se ha obtenido de la matriz* M *trasponiendo dos filas, entonces* $d_{\overline{M}} = -d_M$.

En efecto, supongamos que en la matriz M hay que trasponer las filas que tienen los números de orden i y j. Esto se puede conseguir mediante una cadena de transformaciones siguientes: primero agregamos a la i-ésima fila de la matriz M su j-ésima y obtenemos la matriz M', resultando, por la condición II, $d_{M'} = d_M$. Después restamos de la j-ésima fila de la matriz M' su i-ésima fila y obtenemos la matriz M'' para la que, en virtud de la propiedad (2), se tiene $d_{M''} = d_{M'}$; la j-ésima fila de la matriz M'' se diferenciará de la i-ésima fila de la matriz M en el signo. Agreguemos ahora a la i-ésima fila de la matriz M'' su j-ésima fila. Para la matriz M''' que se obtiene con esta transformación, por la condición II, se tiene $d_{M'''} = d_{M''}$, además, la i-ésima fila de esta matriz coincide con la j-ésima fila de la matriz M. Finalmente, multiplicando la j-ésima fila de la matriz M''', por el número -1, obtendremos la matriz buscada $\overline{M}$. Por consiguiente, en virtud de la condición I, se tiene

$$d_{\overline{M}} = -d_{M'''} = -d_M.$$

(6) *Si la matriz* M' *se ha obtenido de la matriz* M *trasponiendo las filas, y la* i*-ésima fila de la matriz* M', $i = 1, 2, \ldots, n$, es la α_i*-ésima fila de la matriz* M, *entonces,*

$$d_{M'} = \pm d_M;$$

donde el signo más corresponde al caso en que la sustitución

$$\begin{pmatrix} 1 & 2 & \dots & n \\ \alpha_1 & \alpha_2 & \dots & \alpha_n \end{pmatrix}$$

es par, y el signo menos, al caso en que ésta es impar.

En efecto, la matriz M' se puede obtener de la matriz M realizando cierto número de trasposiciones de dos filas, pudiéndose, por consiguiente, aplicar la propiedad (5). Como se sabe por el § 3, la paridad del número de estas trasposiciones determina la paridad de la sustitución indicada anteriormente.

Veamos ahora las matrices $M = (a_{ij})$, $N = (b_{ij})$ y su producto $Q = MN$, en el sentido del § 13. Hallemos el número d_Q. Sabemos que cualquier i-ésima fila de la matriz Q representa una suma de todas las filas de la matriz N, tomadas con los coeficientes a_{i1}, a_{i2}, ..., a_{in}, respectivamente (véase, por ejemplo, el § 14). Sustituyamos todas las filas de la matriz Q por sus expresiones lineales indicadas mediante las filas de la matriz N y apliquemos unas cuantas veces la propiedad (4). Obtendremos que el número d_Q será igual a la suma de los números d_T para todas las matrices posibles T de la forma siguiente: la i-ésima fila de la matriz T, $i = 1, 2, \dots, n$, es igual a la α_i-ésima fila de la matriz N, multiplicada por el número $a_{i\alpha_i}$. Además, en virtud de la propiedad (3), se pueden excluir todas las matrices T para las que existen unos índices i y j, $i \neq j$ tales que $\alpha_i = \alpha_j$; en otras palabras, quedan solamente las matrices T para las que los índices $\alpha_1, \alpha_2, \dots, \alpha_n$ forman una permutación de los números $1, 2, \dots, n$. En virtud de las propiedades I y (6), el número d_T para esta matriz tiene la forma

$$d_T = \pm a_{1\alpha_1} a_{2\alpha_2} \dots a_{n\alpha_n} d_N,$$

donde el signo se determina por la paridad de la sustitución de los índices. De aquí llegamos a la expresión para el número d_Q; después de sacar el factor común d_N de todos los sumandos de la forma d_T, entre paréntesis queda, evidentemente, el determinante $|M|$ de la matriz M en el sentido de la definición constructiva dada en el § 4, es decir,

$$d_Q = |M| \cdot d_N. \qquad (*)$$

Si ahora tomamos por matriz N la matriz unidad E, se tendrá, $Q = M$, de donde, por la propiedad III, $d_N = d_E = 1$, o sea *para cualquier matriz M se cumple la igualdad*

$$d_M = |M|,$$

que es lo que se quería demostrar. **Simultáneamente, sin haber utilizado el teorema de Laplace, queda demostrado de nuevo el teorema del producto de los determinantes:** para esto es suficiente sustituir

los números d_Q y d_N en la igualdad (*) por los determinantes de las matrices correspondientes.

Terminemos estas consideraciones axiomáticas con la demostración de la **independencia** de las condiciones I—III, o sea, con la demostración de que ninguna de estas condiciones es consecuencia de las otras dos.

Para la demostración de la independencia de la condición III, hagamos $d_M = 0$ para cualquier matriz M de n-ésimo orden. Es evidente que las condiciones I y II se cumplen, mientras que la condición III no.

Para la demostración de la independencia de la condición II, supongamos que para cualquier matriz M el número d_M es igual al producto de los elementos situados en la diagonal principal de esta matriz. *Las condiciones I y III se cumplen, mientras que la condición II ya no tiene lugar.* Finalmente, para la demostración de la independencia de la condición I, hagamos $d_M = 1$ para cualquier matriz M. En este caso, las condiciones II y III se cumplirán, mientras que la condición I no.

CAPITULO IV

NUMEROS COMPLEJOS

§ 17. El sistema de los números complejos

En el curso del álgebra elemental varias veces se efectúa un enriquecimiento de las reservas de los números. El alumno que comienza el estudio del álgebra ya conoce por la aritmética los números enteros y quebrados positivos. En esencia, el álgebra comienza con la introducción de los números negativos, o sea, con la formación del primero de los sistemas numéricos fundamentales: del sistema de los *números enteros* que consta de todos los números enteros, positivos y negativos, incluyendo el cero, y del sistema más amplio de los *números racionales*, que consta de todos los números enteros y quebrados, tanto positivos como negativos.

Posteriormente, se efectúa una ampliación del conjunto de los números introduciendo los números irracionales. El sistema, compuesto de todos los números racionales e irracionales, se llama sistema de números *reales*. Ordinariamente, el curso universitario de análisis matemático contiene una construcción rigurosa del sistema de números reales; sin embargo, para nuestra exposición bastan los conocimientos de los números reales que tiene el lector que comienza a estudiar el álgebra superior.

Finalmente, al terminar el curso del álgebra elemental, se amplía el sistema de números reales obteniendo el sistema de *números complejos*. Naturalmente, este sistema de números sigue siendo menos habitual para el lector que el sistema de números reales, a pesar de que posee unas propiedades muy útiles. En el presente capítulo se expondrá de nuevo la teoría de los números complejos con la extensión y plenitud debida.

La introducción de los números complejos es debida al problema siguiente. Es sabido que los números reales no son suficientes para resolver cualquier ecuación cuadrática con coeficientes reales. La ecuación cuadrática más simple, que carece de raíces en el conjunto de los números reales, es

$$x^2 + 1 = 0; \qquad (1)$$

ahora, nos va a interesar solamente esta ecuación. El problema que se nos plantea es: *hay que ampliar el sistema de números reales hasta obtener un sistema tal de números, en el que la ecuación (1) tenga ya raíz.*

Los puntos del plano se tomarán como material de construcción de este nuevo sistema de números. Recordemos que la representación de los números reales por puntos de una línea (basada en que se obtiene una correspondencia biunívoca entre el conjunto de todos los puntos de la recta y el conjunto de todos los números reales, al poner en correspondencia a cada punto de la recta su abscisa; se suponen dados el origen de coordenadas y la unidad de medida) se utiliza sistemáticamente en todas las ramas de las matemáticas y es tan habitual que, ordinariamente, no hacemos distinción alguna entre un número real y el punto que le corresponde.

Por lo tanto, *queremos definir un sistema de números que se representen por todos los puntos del plano.* Hasta ahora, no hemos tenido que sumar o multiplicar los puntos del plano, lo que nos da derecho de elegir la definición de las operaciones con los puntos, preocupándose solamente de que el nuevo sistema de números posea las propiedades que son el motivo de su creación. Al principio, estas definiciones nos parecerán artificiales, sobre todo la del producto. Sin embargo, en el capítulo X se demostrará que **ningunas otras definiciones de las operaciones**, incluso las que a primera vista parecen más naturales, **nos conducirían al objetivo**, que consiste en la construcción de una ampliación del sistema de números reales, para que la ecuación (1) tenga raíz. Allí mismo se demostrará que, *en esta construcción, la sustitución de los puntos del plano por otro material, no nos conduciría a un sistema de números diferente, por sus propiedades algebraicas, del sistema de números complejos que vamos a construir a continuación.*

Supongamos que en el plano se ha elegido un sistema rectangular de coordenadas. Convengamos en designar los puntos del plano con las letras α, β, γ, ... y en representar con la notación (a, b) el punto α de abscisa a y ordenada b, es decir, que apartándonos un poco de lo convenido en la geometría analítica, escribiremos $\alpha = (a, b)$. Dados los puntos $\alpha = (a, b)$ y $\beta = (c, d)$, **llamaremos** *suma* de estos puntos al punto que tiene la abscisa $a + c$ y la ordenada $b + d$, o sea,

$$(a, b) + (c, d) = (a + c,\ b + d); \qquad (2)$$

llamaremos *producto* de los puntos $\alpha = (a, b)$ y $\beta = (c, d)$ al punto de abscisa $ac - bd$ y ordenada $ad + bc$, o sea,

$$(a, b)(c, d) = (ac - bd,\ ad + bc). \qquad (3)$$

De este modo, hemos definido dos operaciones algebraicas en el conjunto de todos los puntos del plano. Demostremos que *estas operaciones poseen todas las propiedades principales que ellas mismas tienen en el sistema de números reales o en el sistema de números racionales; ambas son conmutativas y asociativas, están ligadas por la ley distributiva y para ellas existen las operaciones inversas: la resta y la división (excluyendo la división por cero).*

Las leyes conmutativa y asociativa de la suma son evidentes (o más exatamente, se deducen de las propiedades correspondientes de la suma de los números reales), puesto que al sumar los puntos en el plano, se suman por separado sus abscisas y sus ordenadas. La conmutatividad del producto se basa en que en la definición del producto los puntos α y β gozan de simetría. Las siguientes igualdades:

$$[(a, b)(c, d)](e, f) = (ac - bd, ad + bc)(e, f) =$$
$$= (ace - bde - adf - bcf, acf - bdf + ade + bce),$$
$$(a, b)[(c, d)(e, f)] = (a, b)(ce - df, cf + de) =$$
$$= (ace - adf - bcf - bde, acf + ade + bce - bdf),$$

demuestran que para el producto se cumple la ley asociativa. La ley distributiva se deduce de las igualdades:

$$[(a, b) + (c, d)](e, f) = (a + c, b + d)(e, f) =$$
$$= (ae + ce - bf - df, af + cf + be + de),$$
$$(a, b)(e, f) + (c, d)(e, f) = (ae - bf, af + be) + (ce - df, cf + de) =$$
$$= (ae - bf + ce - df, af + be + cf + de).$$

Veamos la cuestión de las operaciones inversas. Si se han dado los puntos $\alpha = (a, b)$ y $\beta = (c, d)$, su diferencia será un punto (x, y) tal que

$$(c, d) + (x, y) = (a, b).$$

De aquí, en virtud de (2), se deduce que

$$c + x = a, \quad d + y = b.$$

Por lo tanto, la *diferencia* de los puntos $\alpha = (a, b)$ y $\beta = (c, d)$ es el punto

$$\alpha - \beta = (a - c, b - d) \tag{4}$$

y esta diferencia queda definida unívocamente. En particular, el origen de coordenadas (0,0) sirve de *cero*, y el punto *opuesto* al punto $\alpha = (a, b)$ será el punto

$$-\alpha = (-a, -b). \tag{5}$$

Supongamos ahora que se dan los puntos $\alpha = (a, b)$ y $\beta = (c, d)$, y que el punto β es diferente de cero, o sea, que al menos una de las coordenadas c, d no es igual a cero y, por consiguiente, $c^2 + d^2 \neq 0$. El cociente de la división de α por β tiene que ser un punto (x, y) tal que $(c, d)(x, y) = (a, b)$. De aquí, en virtud de (3), se tiene que,

$$cx - dy = a,$$
$$dx + cy = b.$$

Resolviendo este sistema de ecuaciones, obtenemos:

$$x = \frac{ac + bd}{c^2 + d^2}, \quad y = \frac{bc - ad}{c^2 + d^2}.$$

Por lo tanto, para $\beta \neq 0$, el *cociente* $\frac{\alpha}{\beta}$ existe y se determina unívocamente:

$$\frac{\alpha}{\beta} = \left(\frac{ac + bd}{c^2 + d^2}, \frac{bc - ad}{c^2 + d^2}\right). \tag{6}$$

Poniendo aquí $\beta = \alpha$, obtenemos que en nuestra multiplicación de los puntos la *unidad* es el punto $(1, 0)$, situado en el eje de abscisas a la distancia 1 del origen de coordenadas a la derecha. Poniendo luego en (6) $\alpha = 1 = (1, 0)$, obtenemos que, para $\beta \neq 0$, el punto *opuesto* a β es:

$$\beta^{-1} = \left(\frac{c}{c^2 + d^2}, \frac{-d}{c^2 + d^2}\right). \tag{7}$$

Por lo tanto, hemos construido un sistema de números representados por puntos del plano, donde las operaciones con ellos quedan definidas por las fórmulas (2) *y* (3); *éste se denomina sistema de números complejos.*

Demostremos que *este sistema representa una ampliación del sistema de números reales.* Con este fin, veamos los puntos situados en el eje de abscisas, o sea, los puntos de la forma $(a, 0)$; poniendo en correspondencia al *punto* $(a, 0)$, *el número real* a, *obtenemos* evidentemente una correspondencia biunívoca entre el conjunto considerado de puntos y el conjunto de todos los números reales. (Véase la nota del T. en la pág. 25) La aplicación de las fórmulas (2) y (3) a estos puntos proporciona las igualdades

$$(a, 0) + (b, 0) = (a + b, 0),$$
$$(a, 0) \cdot (b, 0) = (ab, 0),$$

o sea, los puntos $(a, 0)$ se suman y se multiplican entre sí, igual que los números reales correspondientes. Por lo tanto, el *conjunto de puntos situados en el eje de abscisas, considerado como una parte del sistema de números complejos, no se diferencia en nada por sus*

propiedades algebraicas del sistema de números reales, representado ordinariamente por puntos de una recta. Esto nos permite no hacer a continuación ninguna distinción entre el punto $(a, 0)$ y el número real a, o sea, que pondremos $(a, 0) = a$. En particular, el cero $(0, 0)$ y la unidad $(1, 0)$ del sistema de números complejos resultan ser los números reales ordinarios 0 y 1.

Tenemos que mostrar ahora que *entre los números complejos está contenida una raíz de la ecuación* (1), es decir, un número cuyo cuadrado sea igual al número real -1. Este será, por ejemplo, el punto $(0, 1)$, o sea, el punto situado en el eje de ordenadas a la distancia 1 del origen de coordenadas, hacia arriba. En efecto, aplicando (3), obtenemos:

$$(0, 1)\cdot(0, 1) = (-1, 0) = -1.$$

Designemos este punto con la letra i, de modo que $i^2 = -1$.

Finalmente, demostremos que *para los números complejos introducidos se puede obtener su expresión ordinaria.* Para esto, hallemos primero el producto del número real b por el punto i.

$$bi = (b, 0)\cdot(0, 1) = (0, b);$$

por consiguiente, éste es el punto que tiene la ordenada b y está situado en el eje de ordenadas; además todos los puntos del eje de ordenadas se representan en forma de productos de éstos. Si ahora (a, b) es un punto arbitrario, en virtud de la igualdad

$$(a, b) = (a, 0) + (0, b),$$

se tiene:

$$(a, b) = a + bi,$$

o sea, que verdaderamente llegamos a la expresión ordinaria de los números complejos; por supuesto, en la expresión $a + bi$, la suma y el producto se deben entender en el sentido de las operaciones definidas en el sistema de números complejos construido.

Una vez introducidos los números complejos, el lector comprobará fácilmente que *todo el contenido de los capítulos precedentes del libro* (la teoría de los determinantes, la teoría de los sistemas de ecuaciones lineales, la teoría de la dependencia lineal de los vectores y la teoría de las operaciones con las matrices) *se generaliza sin restricciones al caso en que se permite el uso de cualesquiera números complejos, y no sólo de los números reales.*

Por último, obsérvese que la construcción expuesta del sistema de números complejos nos lleva a la siguiente pregunta: ¿Se puede definir la suma y el producto de los puntos del espacio de tres dimensiones, de modo que el conjunto de éstos forme un sistema de números que contenga al sistema de números complejos o, al menos, al sistema de números reales? Esta cuestión sale fuera de

los márgenes de nuestro curso y solamente señalaremos que la respuesta es negativa.

Por otra parte, observando que la suma de los números complejos definida anteriormente coincide en su esencia con la suma de vectores en el plano que parten del origen de coordenadas (véase el siguiente párrafo), resulta natural la siguiente pregunta: ¿Es posible definir para ciertos valores de n el producto de vectores del espacio vectorial real de n dimensiones, de modo que éste sea, con respecto a esta multiplicación y a la adición ordinaria de los vectores, un sistema numérico que contenga al sistema de números reales? Se puede demostrar que esto no se puede hacer si se quiere que se *cumplan todas las propiedades de las operaciones que tienen* lugar en los sistemas de números racionales, reales y complejos. *En el espacio de cuatro dimensiones esta construcción es posible* si se prescinde de la conmutatividad de la multiplicación; el sistema de números obtenido se denomina *sistema de cuaterniones*. También es posible una construcción análoga en el espacio de ocho dimensiones, resultando el llamado sistema de *números de Cayley*. Desde luego, en este caso no hay que prescindir solamente de la conmutatividad del producto, sino también de su asociatividad, sustituyendo esta última por otra menos rigurosa.

§ 18. Estudio posterior de los números complejos

De acuerdo a la tradición histórica, al número complejo i lo llamaremos *unidad imaginaria*, y a los números de la forma bi, *números imaginarios puros*, a pesar de que no dudamos de la existencia de ellos, pudiendo señalar los puntos del plano (que están en el eje de ordenadas) que los representan. En la expresión del número complejo α en la forma $\alpha = a + bi$, el número a se denomina *parte real* del número α, y el número bi, *parte imaginaria*. El plano, cuyos puntos se han identificado con los números complejos según el método expuesto en el § 17, se llamará *plano complejo*. El eje de abscisas de este plano se llama *eje real*, puesto que sus puntos representan a los números reales; respectivamente, el eje de ordenadas del plano complejo se *llama eje imaginario*.

La suma, resta, multiplicación y división de los números complejos expresados en la forma $a + bi$, como se deduce de las fórmulas (2), (4), (3), y (6) del párrafo anterior, se efectúan del modo siguiente:

$$(a + bi) + (c + di) = (a + c) + (b + d)\, i;$$
$$(a + bi) + (c - di) = (a - c) + (b - d)\, i;$$
$$(a + bi)(c + di) = (ac - bd) + (ad + bc)\, i;$$
$$\frac{a + bi}{c + di} = \frac{ac + bd}{c^2 + d^2} + \frac{bc - ad}{c^2 + d^2}\, i.$$

Se puede decir que *al sumar los números complejos, se suman por separado sus partes reales y sus partes imaginarias*; para la resta se cumple una regla análoga. Las expresiones verbales de las fórmulas para multiplicar y dividir serían muy complicadas y las omitimos. No hay necesidad de recordar la última de estas fórmulas:

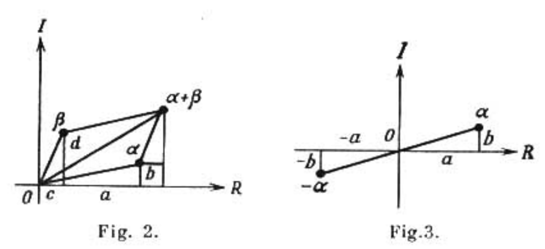


Fig. 2. Fig. 3.

solamente hay que tener en cuenta que ésta se puede deducir multiplicando el numerador y denominador del quebrado dado por un número, que se diferencia del denominador solamente por el signo de la parte imaginaria.

En efecto,

$$\frac{a+bi}{c+di}=\frac{(a+bi)(c-di)}{(c+di)(c-di)}=\frac{(ac+bd)+(bc-ad)i}{c^2+d^2}=\frac{ac+bd}{c^2+d^2}+\frac{bc-ad}{c^2+d^2}i.$$

Ejemplos.

1) $(2+5i)+(1-7i)=(2+1)+(5-7)i=3-2i;$

2) $(3-9i)-(7+i)=(3-7)+(-9-1)i=-4-10i;$

3) $(1+2i)(3-i)=[1\cdot 3-2\cdot(-1)]+[1\cdot(-1)+2\cdot 3]i=5+5i;$

4) $\frac{23+i}{3+i}=\frac{(23+i)(3-i)}{(3+i)(3-i)}=\frac{70-20i}{10}=7-2i.$

La representación de los números complejos por puntos del plano conduce al deseo natural de obtener una interpretación geométrica de las operaciones definidas para los números complejos. Esta es fácil de lograr para la suma. Sean dados los números $\alpha = a+bi$ y $\beta = c+di$. Unamos con segmentos el origen de coordenadas con los puntos (a, b) y (c, d) correspondientes a dichos números, y sobre estos segmentos, como lados, trazemos un paralelogramo (fig. 2). Es evidente que el cuarto vértice de este paralelogramo será el punto $(a+c, b+d)$. Por lo tanto, la *suma de números complejos se efectúa geométricamente por la regla del paralelogramo, o sea por la regla de la suma de vectores que parten del origen de coordenadas.*

El número opuesto al número $\alpha = a + bi$ es el punto del plano complejo que es simétrico al punto α con respecto del origen de coordenadas (fig. 3). De aquí se puede obtener sin dificultad alguna la interpretación geométrica de la resta.

La interpretación geométrica de la multiplicación y división de los números complejos quedará clara solamente después de que introduzcamos una nueva *expresión para los números complejos*. Para la expresión del número α en la forma $\alpha = a + bi$ utilizamos las coordenadas cartesianas del punto correspondiente a este número. Sin embargo, la posición del punto en el plano queda también determinada, si se conocen sus *coordenadas polares: la distancia* r del origen de coordenadas al punto y el ángulo φ que forma la dirección positiva del eje de abscisas con la dirección que va desde el origen de coordenadas hacia este punto (fig. 4).

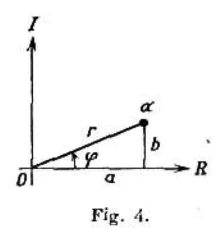


Fig. 4.

El número r es real y no negativo, siendo además igual a cero solamente para el punto 0. Para un número α situado en el eje real, o sea, para un número real, el número r es el valor absoluto de α; por esto, a veces, para *cualquier número complejo α, a r también se le llaman valor absoluto* o *módulo* del número α, representándose por la notación $|\alpha|$.

El ángulo φ se llamará *argumento* del número α y se designará con la notación: arg α*. El ángulo φ puede tomar cualesquiera valores reales, tanto positivos como negativos, teniendo que medir*se los ángulos positivos en dirección contraria a la del movimien*to de las agujas del reloj; sin embargo, si los ángulos se diferencian entre sí en 2π o en un número múltiplo de 2π, sus puntos correspondientes del plano coinciden.

De este modo, el argumento de un número complejo α tiene *infinitos valores, que se diferencian entre sí en números enteros* múltiplos de 2π; por consiguiente, de la igualdad de dos números complejos, representados por sus módulos y sus argumentos, solamente se puede hacer la conclusión de que sus argumentos se diferencian en un número entero múltiplo de 2π, mientras que sus módu*los son iguales. Solamente para el número 0 el argumento es indefi*nido; sin embargo, este número queda completamente determinado por la igualdad: $|0| = 0$.

El argumento del número complejo es una generalización natural del signo del número real. En efecto, el argumento de un número *real positivo es igual a cero, el argumento de un número real nega-*

* No recurrimos a las denominaciones corrientes de las coordenadas polares: radio polar y ángulo polar.

tivo es igual a π; en el eje real, del origen de coordenadas parten solamente dos direcciones, las cuales se pueden distinguir por los símbolos: + y —. En el plano complejo hay infinitas direcciones que parten del punto 0, diferenciándose por el ángulo que forman con la dirección positiva del eje real.

Las coordenadas cartesianas y polares de un punto están ligadas por las relaciones siguientes:

$$a = r\cos\varphi,\quad b = r\operatorname{sen}\varphi, \tag{1}$$

que se cumplen independientemente de la posición del punto en el plano.

De aquí que

$$r = +\sqrt{a^2+b^2}. \tag{2}$$

Apliquemos las fórmulas (1) a un número complejo arbitrario $\alpha = a + bi$:

$$\alpha = a + bi = r\cos\varphi + (r\operatorname{sen}\varphi)\,i,$$

o sea,

$$\alpha = r(\cos\varphi + i\operatorname{sen}\varphi). \tag{3}$$

Recíprocamente, supongamos que el número $\alpha = a + bi$ se expresa en la forma $\alpha = r_0(\cos\varphi_0 + i\operatorname{sen}\varphi_0)$, donde r_0 y φ_0 son unos números reales, siendo $r_0 \geqslant 0$. Entonces, $r_0\cos\varphi_0 = a$, $r_0\operatorname{sen}\varphi_0 = b$, de donde $r_0 = +\sqrt{a^2+b^2}$, y, en virtud de (2), $r_0 = |\alpha|$. De aquí, aplicando (1), obtenemos: $\cos\varphi_0 = \cos\varphi$, $\operatorname{sen}\varphi_0 = \operatorname{sen}\varphi$, o sea, $\varphi_0 = \arg\alpha$. Por lo tanto, *todo número complejo α se expresa unívocamente en la forma* (3), *donde $r = |\alpha|$, $\varphi = \arg\alpha$* (por supuesto, el argumento φ está definido salvo un sumando, múltiplo de 2π). Esta expresión del número α se llama *forma trigonométrica* y se empleará frecuentemente a continuación.

Los números

$$\alpha = 3\left(\cos\frac{\pi}{4} + i\operatorname{sen}\frac{\pi}{4}\right),\quad \beta = \cos\frac{19}{3}\pi + i\operatorname{sen}\frac{19}{3}\pi$$

y

$$\gamma = \sqrt{3}\left[\cos\left(-\frac{\pi}{7}\right) + i\operatorname{sen}\left(-\frac{\pi}{7}\right)\right]$$

están dados en forma trigonométrica; aquí, $|\alpha| = 3$, $|\beta| = 1$, $|\gamma| = \sqrt{3}$; $\arg\alpha = \frac{\pi}{4}$, $\arg\beta = \frac{19}{3}\pi$, $\arg\gamma = -\frac{\pi}{7}$ $\left(\text{o bien, } \arg\beta = \frac{\pi}{3},\ \arg\gamma = \frac{13}{7}\pi\right)$.

Por otra parte, los números complejos

$$\alpha' = (-2)\left(\cos\frac{\pi}{5} + i\operatorname{sen}\frac{\pi}{5}\right),\quad \beta' = 3\left(\cos\frac{2}{3}\pi - i\operatorname{sen}\frac{2}{3}\pi\right),$$

$$\gamma' = 2\left(\cos\frac{\pi}{3} + i\operatorname{sen}\frac{3}{4}\pi\right),\ \delta' = \operatorname{sen}\frac{3}{4}\pi + i\cos\frac{3}{4}\pi$$

ya no están dados en forma trigonométrica, a pesar de que estas expresiones se

parezcan a la expresión (3). Estos números se expresan en forma trigonométrica del modo siguiente:

$$\alpha' = 2\left(\cos\frac{6}{5}\pi + i\ \mathrm{sen}\ \frac{6}{5}\pi\right), \quad \beta' = 3\left(\cos\frac{4}{3}\pi + i\ \mathrm{sen}\ \frac{4}{3}\pi\right),$$

$$\delta' = \cos\frac{7}{4}\pi + i\ \mathrm{sen}\ \frac{7}{4}\pi.$$

La determinación de la forma trigonométrica del número γ' es engorrosa, como casi siempre ocurre al pasar de la expresión ordinaria del número complejo a la trigonométrica y viceversa: salvo unos pocos casos, dados los valores numéricos del seno y del coseno, resulta imposible hallar el valor exacto del ángulo, dado el ángulo, resulta imposible hallar los valores *exactos* de su seno y coseno.

Supongamos que los números complejos α y β se dan en su forma trigonométrica: $\alpha = r\ (\cos\varphi + i\ \mathrm{sen}\ \varphi)$, $\beta = r'\ (\cos\varphi' + i\ \mathrm{sen}\ \varphi')$. Multipliquemos estos números:

$$\alpha\beta = [r(\cos\varphi + i\ \mathrm{sen}\ \varphi)]\cdot[r'(\cos\varphi' + i\ \mathrm{sen}\ \varphi')] =$$
$$= rr'(\cos\varphi\cos\varphi' + i\cos\varphi\ \mathrm{sen}\ \varphi' + i\ \mathrm{sen}\ \varphi\cos\varphi' - \mathrm{sen}\ \varphi\ \mathrm{sen}\ \varphi'),$$

o sea,

$$\alpha\beta = rr'\,[\cos(\varphi+\varphi') + i\ \mathrm{sen}\ (\varphi+\varphi')]. \tag{4}$$

Hemos obtenido la expresión del producto en la forma trigonométrica, de donde $|\alpha\beta| = rr'$, o sea,

$$|\alpha\beta| = |\alpha|\,|\beta|, \tag{5}$$

es decir, el *módulo del producto de números complejos es igual al producto de los módulos de los factores*; por otra parte, $\arg(\alpha\beta) = \varphi + \varphi'$, o sea,

$$\arg(\alpha\beta) = \arg\alpha + \arg\beta, \tag{6}$$

es decir, el *argumento del producto de números complejos es igual a la suma de los argumentos de los factores**. Claro que estas reglas *se generalizan para cualquier número finito de factores. En el caso* de números reales, la fórmula (5) proporciona la conocida propiedad de los valores absolutos de estos números, mientras que la fórmula (6), como fácilmente se comprueba, se convierte en la regla *de los signos de la multiplicación de los números reales.*

La división también goza de propiedades análogas. En efecto, sea $\alpha = r\ (\cos\varphi + i\ \mathrm{sen}\ \varphi)$, $\beta = r'\ (\cos\varphi' + i\ \mathrm{sen}\ \varphi')$, siendo $\beta \neq 0$, o sea, $r' \neq 0$. Entonces,

$$\frac{\alpha}{\beta} = \frac{r(\cos\varphi + i\ \mathrm{sen}\ \varphi)}{r'(\cos\varphi' + i\ \mathrm{sen}\ \varphi')} = \frac{r(\cos\varphi + i\ \mathrm{sen}\ \varphi)(\cos\varphi' - i\ \mathrm{sen}\ \varphi')}{r'(\cos^2\varphi' + \mathrm{sen}^2\varphi')} =$$
$$= \frac{r}{r'}(\cos\varphi\cos\varphi' + i\ \mathrm{sen}\ \varphi\cos\varphi' - i\cos\varphi\ \mathrm{sen}\ \varphi' + \mathrm{sen}\ \varphi\ \mathrm{sen}\ \varphi'),$$

* Subrayemos que aquí la igualdad se entiende salvo un sumando, múltiplo de 2π.

o sea,

$$\frac{\alpha}{\beta}=\frac{r}{r'}\,[\cos(\varphi-\varphi')+i\,\text{sen}\,(\varphi-\varphi')]. \tag{7}$$

De aquí se deduce que $\left|\frac{\alpha}{\beta}\right|=\frac{r}{r'}$, o bien,

$$\left|\frac{\alpha}{\beta}\right|=\frac{|\alpha|}{|\beta|}, \tag{8}$$

es decir, el *módulo del cociente de dos números complejos* ***es igual al*** *módulo del dividendo dividido por el módulo del divisor*; por otra parte, $\arg\left(\frac{\alpha}{\beta}\right)=\varphi-\varphi'$, o sea,

$$\arg\left(\frac{\alpha}{\beta}\right)=\arg\alpha-\arg\beta \tag{9}$$

es decir, el *argumento del cociente de dos números complejos se obtiene restando el argumento del divisor del argumento del dividendo.*

El significado geométrico del producto y del cociente se aclara ahora sin dificultad. En efecto, en virtud de las fórmulas (5)

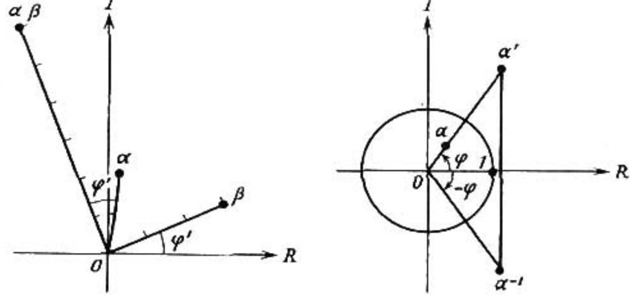


Fig. 5. Fig. 6.

y (6), para obtener el punto que representa el producto del número α por el número $\beta=r'(\cos\varphi'+i\,\text{sen}\,\varphi')$, hay que hacer girar al vector que va de 0 a α (fig. 5) un ángulo $\varphi'=\arg\beta$ en dirección contraria a la del movimiento de las agujas del reloj, y después hay que alargar este vector $r'=|\beta|$ veces (si $0\leqslant r'<1$, esto no será un alargamiento, sino una contracción). Por otra parte, de (7) se deduce que para $\alpha=r\,(\cos\,\varphi+i\,\,\text{sen}\,\,\varphi)\neq 0$, se tiene,

$$\alpha^{-1}=r^{-1}\,[\cos(-\varphi)+i\,\text{sen}\,(-\varphi)], \tag{10}$$

o sea, $|\alpha|^{-1}=|\alpha^{-1}|$, $\arg(\alpha^{-1})=-\arg\alpha$. Por lo tanto, para obtener el punto α^{-1} hay que pasar del punto α al punto α', situado

en la misma semirrecta que parte del cero y que pasa por el punto α, a la distancia r^{-1} del cero (fig 6)*, y después hay que pasar al punto simétrico a α' con respecto al eje real.

La suma y la diferencia de números complejos, dados en forma trigonométrica, no se pueden expresar por fórmulas semejantes a las fórmulas (4) y (7). Sin embargo, para el módulo de la suma se cumplen las *importantes desigualdades:*

$$|\alpha|-|\beta|\leqslant|\alpha+\beta|\leqslant|\alpha|+|\beta|. \tag{11}$$

es decir, el *módulo de la suma de dos números complejos es menor o igual a la suma de los módulos de los sumandos, pero es mayor o igual a la diferencia de estos módulos.* Las desigualdades (11) se deducen del conocido teorema de geometría elemental sobre los lados del triángulo, puesto que como se sabe, $|\alpha+\beta|$ es igual a la diagonal del paralelogramo de lados $|\alpha|$ y $|\beta|$. Si los puntos α, β y 0 están situados en una recta, se necesita un estudio especial; esto lo dejamos a cuenta del lector. Solamente en este caso se cumple el signo de igualdad en las fórmulas (11).

Como $\alpha-\beta=\alpha+(-\beta)$ y

$$|-\beta|=|\beta| \tag{12}$$

(esta igualdad es consecuencia de la interpretación geométrica del número $-\beta$), de (11) se deducen también las desigualdades

$$|\alpha|-|\beta|\leqslant|\alpha-\beta|\leqslant|\alpha|+|\beta|,^{**} \tag{13}$$

es decir, para el módulo de la diferencia se cumplen también las mismas desigualdades que para el módulo de la uma.

Las desigualdades (11) se podrían obtener también del modo siguiente: Sea $\alpha = r\ (\cos\varphi + i \operatorname{sen}\varphi)$, $\beta = r'\ (\cos\varphi' + i \operatorname{sen}\varphi')$; supongamos que la forma trigonométrica del número $\alpha+\beta$ es: $\alpha+\beta = R\ (\cos\psi + i \operatorname{sen}\psi)$. Sumando por separado las partes reales y las partes imaginarias, obtenemos:

$$r\cos\varphi + r'\cos\varphi' = R\cos\psi,$$

$$r\operatorname{sen}\varphi + r'\operatorname{sen}\varphi' = R\operatorname{sen}\psi;$$

multiplicando ambos miembros de la primera igualdad por $\cos\psi$, ambos miembros de la segunda por $\operatorname{sen}\psi$ y sumando, obtenemos: $r\ (\cos\varphi\cos\psi +$

* La igualdad $|\alpha'|=|\alpha|$ se cumple cuando, y sólo cuando $|\alpha|=1$, o sea, si el punto α está situado en la circunferencia del *círculo unidad*. Si α está situado dentro del círculo unidad, α' estará situado fuera de él, y viceversa, obteniendo de este modo una correspondencia biunívoca entre todos los puntos del plano complejo, situados fuera del círculo unidad, y todos los puntos, situados dentro de este círculo y *diferentes de cero*.

** Por consiguiente, se cumple también la desigualdad

$$\big||\alpha|-|\beta|\big|\leqslant|\alpha-\beta|,$$

que se aplicará en el § 23. (*Nota de T.*)

$+ \operatorname{sen} \varphi \operatorname{sen} \psi) + r' (\cos \varphi' \cos \psi + \operatorname{sen} \varphi' \operatorname{sen} \psi) = R (\cos^2 \psi + \operatorname{sen}^2 \psi),$
o sea,

$$r \cos (\varphi - \psi) + r' \cos (\varphi' - \psi) = R.$$

Como el coseno nunca es mayor que la unidad, de aquí se deduce la desigualdad $r + r' \geqslant R$, o sea, $|\alpha| + |\beta| \geqslant |\alpha + \beta|$. Por otra parte, $\alpha = (\alpha + \beta) - \beta = (\alpha + \beta) + (-\beta)$. De aquí, por lo demostrado y en virtud de (12),

$$|\alpha| \leqslant |\alpha + \beta| + |-\beta| = |\alpha + \beta| + |\beta|,$$

de donde $|\alpha| - |\beta| \leqslant |\alpha + \beta|$.

Es menester observar que los conceptos «mayor» y «menor» no se pueden definir racionalmente para los números complejos, puesto que éstos, a diferencia de los números reales, no se sitúan en una recta, cuyos puntos están ordenados de un modo natural, sino en un plano. Por esto, **los números complejos (no nos referimos a sus módulos) no se pueden unir nunca con el signo de desigualdad.**

Números conjugados. Sea dado un número complejo $\alpha = a + bi$. El número $a - bi$, que se diferencia de α solamente en el signo de la parte imaginaria, se llama número *conjugado* de α y se designa por $\bar{\alpha}$.

Recordemos, que al estudiar la división de los números complejos recurríamos a los números conjugados, a pesar de que no habíamos introducido esta denominación.

Fig. 7.

Es evidente que el número conjugado de $\bar{\alpha}$ es α, es decir, se puede hablar de pares de números conjugados. Los números reales, y solamente éstos, son conjugados consigo mismos.

Geométricamente, los números conjugados son puntos simétricos entre sí con respecto al eje real (fig. 7). De aquí se deducen las igualdades

$$|\bar{\alpha}| = |\alpha|, \arg \bar{\alpha} = -\arg \alpha. \tag{14}$$

La suma y el producto de números complejos conjugados son números reales. En efecto,

$$\left.\begin{aligned} \alpha + \bar{\alpha} &= 2\alpha, \\ \alpha\bar{\alpha} &= a^2 + b^2 = |\alpha|^2. \end{aligned}\right\} \tag{15}$$

La última igualdad muestra que el número $\alpha\bar{\alpha}$ es, incluso, positivo para $\alpha \neq 0$. En el § 24 se verá un teorema que muestra que la propiedad que acabamos de demostrar de los números conjugados es característica para éstos.

La igualdad

$$(a-bi)+(c-di)=(a+c)-(b+d)i$$

muestra que el *número conjugado de la suma de dos números es igual a la suma de los números conjugados con los sumandos*:

$$\overline{\alpha+\beta}=\bar{\alpha}+\bar{\beta}. \tag{16}$$

Análogamente, de la igualdad

$$(a-bi)(c-di)=(ac-bd)-(ad+bc)i$$

resulta que el *número conjugado con producto es igual al producto de los números conjugados con los factores*:

$$\overline{\alpha\beta}=\bar{\alpha}\cdot\bar{\beta}. \tag{17}$$

Una comprobación directa muestra que se verifican también las igualdades

$$\overline{\alpha-\beta}=\bar{\alpha}-\bar{\beta}, \tag{18}$$

$$\overline{\left(\frac{\alpha}{\beta}\right)}=\frac{\bar{\alpha}}{\bar{\beta}}. \tag{19}$$

Demostremos también la siguiente proposición: *si el número α se expresa de cierto modo por los números complejos $\beta_1, \beta_2, \ldots, \beta_n$ mediante la suma, el producto, la resta y la división, entonces, al sustituir en esta expresión todos los números β_k por sus conjugados, se obtiene el número conjugado de α*; en particular, si el número α es real, éste no se altera al sustituir todos los números complejos β_k por sus conjugados.

Esta proposición la demostraremos por inducción sobre n, puesto que para $n=2$ ésta se deduce de las fórmulas (16)—(19).

Supongamos que el número α se expresa por los números $\beta_1, \beta_2, \ldots, \beta_n$, que no son necesariamente diferentes. En esta expresión hay un orden determinado de aplicación de las operaciones de sumar, multiplicar, restar y dividir. El último acto consistirá en la aplicación de una de estas operaciones a un número γ_1, expresado mediante los números $\beta_1, \beta_2, \ldots, \beta_k$, donde $1 \leqslant k \leqslant n-1$, y a un número γ_2, expresado mediante los números $\beta_{k+1}, \ldots, \beta_n$. Por la hipótesis de inducción, la sustitución de los números $\beta_1, \beta_2, \ldots \beta_k$ por los conjugados implica el cambio del número γ_1 por $\overline{\gamma_1}$, y la sustitución de los números $\beta_{k+1}, \beta_{k+2}, \ldots, \beta_n$ por los conjugados, el cambio del número γ_2 por $\overline{\gamma_2}$. Pero, según una de las fórmulas (16)—(19), el cambio de γ_1 y γ_2 por $\overline{\gamma_1}$ y $\overline{\gamma_2}$ convierte al número α en el número $\bar{\alpha}$.

§ 19. Extracción de la raíz de los números complejos

Estudiemos el problema de la elevación de los números complejos a una potencia y de la extracción de una raíz. Para elevar el número $\alpha = a + bi$ a una potencia entera y positiva n, es suficiente aplicar la fórmula del binomio de Newton a la expresión $(a + bi)^n$ (esta fórmula subsiste también para los números complejos, puesto que su demostración se basa solamente en la ley distributiva), y después, las igualdades: $i^2 = -1$, $i^3 = -i$, $i^4 = 1$; en general

$$i^{4k} = 1, \quad i^{4k+1} = i, \quad i^{4k+2} = -1, \quad i^{4k+3} = -i.$$

Si el número α está dado en forma trigonométrica, entonces, siendo n entero y positivo, de la fórmula (4) del párrafo anterior resulta la fórmula siguiente, llamada *fórmula de Moivre*:

$$[r(\cos\varphi + i \operatorname{sen}\varphi)]^n = r^n(\cos n\varphi + i \operatorname{sen} n\varphi), \tag{1}$$

o sea, que *al elevar un número complejo a una potencia, se eleva el módulo a esta potencia y se multiplica el argumento por el exponente de la potencia.* La fórmula (1) es válida también para los exponentes enteros negativos. En efecto, en virtud de la igualdad $\alpha^{-n} = (\alpha^{-1})^n$, es suficiente aplicar la fórmula de Moivre al número α^{-1}, cuya forma trigonométrica viene dada por la fórmula (10) del párrafo anterior.

Ejemplos.

1) $i^{37} = i$, $i^{122} = -1$;

2) $(2 + 5i)^3 = 2^3 + 3\cdot 2^2\cdot 5i + 3\cdot 2\cdot 5^2 i^2 + 5^3 i^3 =$
$= 8 + 60i - 150 - 125i = -142 - 65i$;

3) $\left[\sqrt{2}\left(\cos\frac{\pi}{4} + i \operatorname{sen}\frac{\pi}{4}\right)\right]^4 = (\sqrt{2})^4(\cos\pi + i \operatorname{sen}\pi) = -4$;

4) $\left[3\left(\cos\frac{\pi}{5} + i \operatorname{sen}\frac{\pi}{5}\right)\right]^{-3} =$

$$= 3^{-3}\left[\cos\left(-\frac{3}{5}\pi\right) + i \operatorname{sen}\left(-\frac{3}{5}\pi\right)\right] = \frac{1}{27}\left(\cos\frac{7}{5}\pi + i \operatorname{sen}\frac{7}{5}\pi\right).$$

De la igualdad

$$(\cos\varphi + i \operatorname{sen}\varphi)^n = \cos n\varphi + i \operatorname{sen} n\varphi,$$

que representa un caso particular de la fórmula de Moivre, fácilmente se obtienen las fórmulas para el seno y el coseno de un ángulo múltiplo. En efecto, aplicando la fórmula del binomio de Newton al primer miembro de esta igualdad e igualando por separado las

partes reales e imaginarias de ambos miembros, se tiene:

$$\cos n\varphi = \cos^n \varphi - \binom{n}{2} \cos^{n-2}\varphi \cdot \operatorname{sen}^2 \varphi + \binom{n}{4} \cos^{n-4}\varphi \cdot \operatorname{sen}^4 \varphi - \ldots,$$

$$\operatorname{sen} n\varphi = \binom{n}{1} \cos^{n-1} \varphi \cdot \operatorname{sen} \varphi - \binom{n}{3} \cos^{n-3} \varphi \cdot \operatorname{sen}^3 \varphi +$$
$$+ \binom{n}{5} \cos^{n-5}\varphi \cdot \operatorname{sen}^5\varphi - \ldots;$$

aquí $\binom{n}{k}$ es la notación ordinaria del coeficiente binomial

$$\binom{n}{k} = \frac{n(n-1)(n-2)\ldots(n-k+1)}{1\cdot 2\cdot 3\ldots k}.$$

Para $n=2$, se tienen las conocidas fórmulas

$$\cos 2\varphi = \cos^2 \varphi - \operatorname{sen}^2 \varphi,$$
$$\operatorname{sen} 2\varphi = 2 \cos \varphi \operatorname{sen} \varphi;$$

para $n=3$,

$$\cos 3\varphi = \cos^3 \varphi - 3 \cos \varphi \operatorname{sen}^2 \varphi,$$
$$\operatorname{sen} 3\varphi = 3 \cos^2 \varphi \operatorname{sen} \varphi - \operatorname{sen}^3 \varphi.$$

La extracción de la raíz de los números complejos es mucho más complicada. Comencemos por la extracción de la raíz cuadrada del número $\alpha = a + bi$. Todavía no sabemos si existe un número complejo cuyo cuadrado sea igual a α. Suponiendo que tal número existe, por ejemplo, $u + vi$ y empleando la notación ordinaria, se puede escribir:

$$\sqrt{a+bi} = u + vi.$$

De la igualdad

$$(u+vi)^2 = a + bi$$

resulta,

$$\left.\begin{aligned} u^2 - v^2 &= a, \\ 2uv &= b. \end{aligned}\right\} \quad (2)$$

Elevando al cuadrado ambos miembros de las igualdades (2) y sumándolas después, se tiene

$$(u^2 - v^2)^2 + 4u^2v^2 = (u^2 + v^2)^2 = a^2 + b^2,$$

de donde

$$u^2 + v^2 = +\sqrt{a^2 + b^2};$$

se toma el signo más, porque los números u y v son reales, debido a lo cual, el primer miembro de la igualdad es positivo. De esta igual-

dad y de la primera de las igualdades (2), resulta:

$$u^2 = \frac{1}{2}(a + \sqrt{a^2+b^2}),$$

$$v^2 = \frac{1}{2}(-a + \sqrt{a^2+b^2}).$$

Extrayendo las raíces cuadradas se obtienen dos valores para u, que se diferencian en el signo, y también dos valores para v. Todos estos valores son reales, puesto que para cualesquiera a y b, las raíces cuadradas se extraen de números positivos. Los valores obtenidos de u y v no se pueden combinar entre sí de modo arbitrario, puesto que, en virtud de la segunda de las igualdades (2), el signo del producto uv tiene que coincidir con el signo de b. Resultan, pues, dos combinaciones posibles de los valores de u y v, o sea, dos números de la forma $u + vi$, que pueden servir de valores de la raíz cuadrada del número α; estos números se diferencian entre sí en el signo. Una prueba elemental, aunque complicada (elevando al cuadrado los números obtenidos, una vez cuando $b > 0$, y otra vez cuando $b < 0$), muestra que los números obtenidos son, verdaderamente, valores de la raíz cuadrada del número α. Por lo tanto, *siempre es posible la extracción de la raíz cuadrada de un número complejo, proporcionando ésta dos valores que se diferencian entre sí en el signo.*

En particular, ahora resulta posible la extracción de la raíz cuadrada de un número real negativo, siendo los valores de esta raíz números imaginarios puros. En efecto, si $a < 0$ y $b = 0$, entonces $\sqrt{a^2 + b^2} = -a$, puesto que esta raíz tiene que ser positiva, por lo cual, $u^2 = \frac{1}{2}(a - a) = 0$, o sea, $u = 0$, así que $\sqrt{a} = \pm vi$.

Ejemplo. Sea $\alpha = 21 - 20i$. Entonces $\sqrt{a^2 + b^2} = \sqrt{441 + 400} = 29$. Por consiguiente, $u^2 = \frac{1}{2}(21 + 29) = 25$, $v^2 = \frac{1}{2}(-21 + 29) = 4$, de donde $u = \pm 5$, $v = \pm 2$. Los signos de u y v tienen que ser diferentes, puesto que b es negativo; en consecuencia,

$$\sqrt{21 - 20i} = \pm (5 - 2i).$$

Las pruebas de extracción de raíces de grado más elevado de los números complejos, dados en la forma $a + bi$, chocan con dificultades insuperables. Así, pues, si quisiéramos hallar con el mismo método la raíz cúbica del número $a + bi$, tendríamos que resolver una ecuación cúbica auxiliar, cosa que por el momento no sabemos hacer y que, como ya veremos en el § 38, requiere a su vez la extracción de la raíz cúbica de números complejos. Por otra parte, la forma trigonométrica se adapta perfectamente para la extracción de las raíces de cualquier grado, con cuya aplicación se resuelve totalmente este problem.

Supongamos que se necesita extraer la raíz n-ésima del número $\alpha = r(\cos\varphi + i \operatorname{sen}\varphi)$. Supongamos también que ésta se puede hallar, *resultando* $\rho(\cos\theta + i\operatorname{sen}\theta)$, *de modo que*,

$$[\rho(\cos\theta + i\operatorname{sen}\theta)]^n = r(\cos\varphi + i\operatorname{sen}\varphi). \qquad (3)$$

Según la fórmula de Moivre, $\rho^n = r$, *o sea*, $\rho = \sqrt[n]{r}$, *donde en* el segundo miembro figura el valor positivo, unívocamente determinado, de la raíz n-ésima del número real positivo r. Por otra parte, el argumento del primer miembro de la igualdad (3) es $n\theta$. Sin embargo, no se puede afirmar que $n\theta$ es igual a φ, porque, en realidad, éstos pueden diferir en un sumando que es múltiplo entero del número 2π. En consecuencia, $n\theta = \varphi + 2k\pi$, donde k es entero y

$$\theta = \frac{\varphi + 2k\pi}{n}.$$

Recíprocamente, tomando el número $\sqrt[n]{r}\left(\cos\frac{\varphi+2k\pi}{n} + i\operatorname{sen}\frac{\varphi+2k\pi}{n}\right)$, se tiene que, **para cualquier** k **entero, positivo o negativo,** la n-ésima potencia de este número es igual a α. Por lo tanto

$$\sqrt[n]{r(\cos\varphi + i\operatorname{sen}\varphi)} = \sqrt[n]{r}\left(\cos\frac{\varphi+2k\pi}{n} + i\operatorname{sen}\frac{\varphi+2k\pi}{n}\right). \qquad (4)$$

Dando a k diversos valores, *no siempre se obtienen diversos* valores de la raíz buscada. En efecto, para

$$k = 0, 1, 2, \ldots, n-1 \qquad (5)$$

se obtienen n valores distintos de la raíz, puesto que el aumento de k en una unidad ocasiona un aumento del argumentos en $\frac{2\pi}{n}$. Supongamos ahora que k es arbitrario. Si $k = nq + r$, $0 \leqslant +r \leqslant n-1$, se tiene,

$$\frac{\varphi+2k\pi}{n} = \frac{\varphi+2(nq+r)\pi}{n} = \frac{\varphi+2r\pi}{n} + 2q\pi,$$

es decir, el valor del argumento para nuestro k difiere del valor del argumento para $k = r$ en un número múltiplo de 2π; por consiguiente, se obtiene el mismo valor de la raíz que resulta para el valor de k igual a r, incluido en el sistema (5).

Por lo tanto, *siempre es posible la extracción de la raíz n-ésima de un número complejo* α, *resultando n valores distintos. Todos los valores de la raíz n-ésima están situados en una circunferencia de radio* $\sqrt[n]{|\alpha|}$ *con el centro en el cero, dividiendo a ésta en n partes iguales.*

En particular, la raíz n-ésima de un número real a tiene también n valores distintos; entre éstos puede haber dos reales, uno *o ninguno, dependiendo del signo de* a *y de la paridad de* n.

Ejemplos.

1) $$\beta=\sqrt[3]{2\left(\cos\frac{3}{4}\pi+i\operatorname{sen}\frac{3}{4}\pi\right)}=$$

$$=\sqrt[3]{2}\left(\cos\frac{\frac{3}{4}\pi+2k\pi}{3}+i\operatorname{sen}\frac{\frac{3}{4}\pi+2k\pi}{3}\right);$$

$$k=0:\ \beta_0=\sqrt[3]{2}\left(\cos\frac{\pi}{4}+i\operatorname{sen}\frac{\pi}{4}\right);$$

$$k=1:\ \beta_1=\sqrt[3]{2}\left(\cos\frac{11}{12}\pi+i\operatorname{sen}\frac{11}{12}\pi\right);$$

$$k=2:\ \beta_2=\sqrt[3]{2}\left(\cos\frac{19}{12}\pi+i\operatorname{sen}\frac{19}{12}\pi\right).$$

2) $$\beta=\sqrt{i}=\sqrt{\cos\frac{\pi}{2}+i\operatorname{sen}\frac{\pi}{2}}=$$

$$=\cos\frac{\frac{\pi}{2}+2k\pi}{2}+i\operatorname{sen}\frac{\frac{\pi}{2}+2k\pi}{2};$$

$$\beta_0=\cos\frac{\pi}{4}+i\operatorname{sen}\frac{\pi}{4}=\frac{\sqrt{2}}{2}+i\frac{\sqrt{2}}{2};$$

$$\beta_1=\cos\frac{5}{4}\pi+i\operatorname{sen}\frac{5}{4}\pi=-\beta_0.$$

3) $$\beta=\sqrt[3]{-8}=\sqrt[3]{8(\cos\pi+i\operatorname{sen}\pi)}=$$

$$=2\left(\cos\frac{\pi+2k\pi}{3}+i\operatorname{sen}\frac{\pi+2k\pi}{3}\right);$$

$$\beta_0=2\left(\cos\frac{\pi}{3}+i\operatorname{sen}\frac{\pi}{3}\right)=1+i\sqrt{3};$$

$$\beta_1=2(\cos\pi+i\operatorname{sen}\pi)=-2;$$

$$\beta_2=2\left(\cos\frac{5\pi}{3}+i\operatorname{sen}\frac{5\pi}{3}\right)=1-i\sqrt{3}.$$

Raíces de la unidad. El caso de la extracción de la raíz n-ésima del número 1 es de particular importancia. Esta raíz tiene n valores y, en virtud de la igualdad $1 = \cos 0 + i \operatorname{sen} 0$ y de la fórmula (4), todos ellos, o, como nos expresaremos, todas las *raíces n-ésimas de la unidad*, vienen dadas por la fórmula

$$\sqrt[n]{1}=\cos\frac{2k\pi}{n}+i\operatorname{sen}\frac{2k\pi}{n};\ k=0,1,\ldots,n-1. \tag{6}$$

Los valores reales de la raíz n-ésima de la unidad se obtienen de la fórmula (6) para los valores $k=0$ y $\frac{n}{2}$, si n es par, y para $k=0$, si n es impar. En el plano complejo las raíces n-ésimas de la unidad

están situadas en la circunferencia del círculo unidad y la dividen en n arcos iguales; el número 1 es uno de los puntos de división. De esto se deduce que las raíces n-ésimas de la unidad que no son reales están situadas simétricamente con respecto al eje real, es decir, son conjugadas entre sí.

La raíz cuadrada de la unidad tiene dos valores: 1 y -1. La raíz cuártica de la unidad tiene cuatro valores: 1, -1, i y $-i$. Para lo futuro es conveniente recordar los valores de la *raíz cúbica de la unidad*. En virtud de (6), éstos son los números $\cos\frac{2k\pi}{3}+$ $+i\operatorname{sen}\frac{2k\pi}{3}$, donde $k=0, 1, 2$, o sea, además de la unidad, se tienen los números conjugados entre sí:

$$\left.\begin{aligned}\varepsilon_1=\cos\frac{2\pi}{3}+i\operatorname{sen}\frac{2\pi}{3}=-\frac{1}{2}+i\frac{\sqrt{3}}{2},\\ \varepsilon_2=\cos\frac{4\pi}{3}+i\operatorname{sen}\frac{4\pi}{3}=-\frac{1}{2}-i\frac{\sqrt{3}}{2}.\end{aligned}\right\}\qquad(7)$$

Todos los valores de la raíz n-ésima del número complejo α *se pueden obtener multiplicando uno de estos valores por todas las raíces n-ésimas de la unidad.* En efecto, sea β uno de los valores de la raíz n-ésima del número α, o sea, que $\beta^n=\alpha$, y sea ε un valor arbitrario de la raíz n-ésima de la unidad, o sea, que $\varepsilon^n=1$. Entonces, $(\beta\varepsilon)^n=\beta^n\varepsilon^n=\alpha$, es decir, $\beta\varepsilon$ también es uno de los valores de $\sqrt[n]{\alpha}$. Multiplicando β por cada una de las raíces n-ésimas de la unidad, obtenemos n valores diferentes de la raíz n-ésima del número α, o sea, todos los valores de esta raíz.

Ejemplos. 1) Uno de los valores de la raíz cúbica de -8 es -2. En virtud de (7), los otros dos serán los números $-2\varepsilon_1=1-i\sqrt{3}$ y $-2\varepsilon_2=1+i\sqrt{3}$ (véase el ejemplo anterior 3). 2) $\sqrt[4]{81}$ tiene cuatro valores: 3, -3, $3i$, $-3i$.

El producto de dos raíces n-ésimas de la unidad también es una raíz n-ésima de la unidad. En efecto, si $\varepsilon^n=1$ y $\eta^n=1$, se tiene, $(\varepsilon\eta)^n=$ $=\varepsilon^n\eta^n=1$. Por otra parte, el *número recíproco de la raíz n-ésima de la unidad también es una raíz n-ésima de la unidad.* En efecto, sea $\varepsilon^n=1$. Entonces, de la igualdad $\varepsilon^n.\ \varepsilon^{-1}=1$ resulta, $\varepsilon^n(\varepsilon^{-1})^n=$ $=1$, o sea, $(\varepsilon^{-1})^n=1$. En general, *toda potencia de la raíz n-ésima de la unidad también es una raíz n-ésima de la unidad.*

Toda raíz k-ésima de la unidad también es raíz l-ésima de la unidad para cualquier l, múltiplo de k. De esto se deduce que, considerando todo el conjunto de las raíces n-ésimas de la unidad, algunas de estas raíces también son raíces n'-ésimas de la unidad para ciertos n', divisores del número n. Sin embargo, para todo n existen raíces n-ésimas de la unidad que no son raíces de la unidad de orden menor. Estas se llaman *raíces primitivas* n-ésimas de la

unidad. Su existencia se deduce de la fórmula (6): si se designa con ε_k el valor de la raíz que corresponde al valor dado de k (de modo que $\varepsilon_0 = 1$), en virtud de la fórmula de Moivre (1), se tiene:

$$\varepsilon_1^k = \varepsilon_k.$$

Por consiguiente, ninguna potencia del número ε_1, menor que la n-ésima, será igual a 1, o sea, el número $\varepsilon_1 = \cos\frac{2\pi}{n} + i \operatorname{sen}\frac{2\pi}{n}$ es una raíz primitiva.

La raíz n-ésima de la unidad ε es primitiva cuando, y sólo cuando, sus potencias ε^k, $k = 0, 1, \ldots, n-1$, *son diferentes, es decir, si con ellas se agotan todas las raíces n-ésimas de la unidad.*

En efecto, si todas las potencias indicadas del número ε son diferentes, es evidente que éste es raíz primitiva n-ésima de la unidad. Si, por el contrario, $\varepsilon^k = \varepsilon^l$ para $0 \leqslant k < l \leqslant n-1$, entonces, $\varepsilon^{l-k} = 1$, y en virtud de las desigualdades $1 \leqslant l - k \leqslant n-1$, la raíz ε no será primitiva.

En el caso general, el número ε_1 hallado anteriormente no será la única raíz primitiva n-ésima de la unidad. Para hallar todas estas raíces se aplica el teorema siguiente:

Si ε es una raíz primitiva n-ésima de la unidad, el número ε^k *es una raíz primitiva n-ésima de la unidad cuando, y sólo cuando, k es primo con n.*

En efecto, sea d el máximo común divisor de los números k y n. Si $d > 1$ y $k = dk'$, $n = dn'$, entonces,

$$(\varepsilon^k)^{n'} = \varepsilon^{kn'} = \varepsilon^{k'n} = (\varepsilon^n)^{k'} = 1,$$

o sea, la raíz ε^k resulta raíz n'-ésima de la unidad.

Por otra parte, supongamos que $d = 1$ y que el número ε^k es una raíz m-ésima de la unidad, $1 \leqslant m < n$. Por lo tanto,

$$(\varepsilon^k)^m = \varepsilon^{km} = 1.$$

Como el número ε es una raíz **primitiva** n-ésima de la unidad, lo que implica que pueden ser iguales a la unidad solamente las potencias del mismo cuyos exponentes sean múltiplos de n, el número km es múltiplo de n. Sin embargo, como $1 \leqslant m < n$, resulta que los números k y n no pueden ser primos entre sí, lo que contradice a la hipótesis. Por lo tanto, el número de raíces primitivas n-ésimas de la unidad es igual al número de enteros positivos k, menores de n y primos con k. La expresión de este número que, generalmente, se designa mediante $\varphi(n)$, se puede hallar en cualquier tratado sobre la teoría de los números. Si p es un número primo, todas las raíces p-ésimas de la unidad son primitivas, a excepción de la unidad misma. Por otra parte, entre las raíces cuárticas de la unidad son primitivas i y $-i$, pero no 1 y -1.

CAPITULO V

LOS POLINOMIOS Y SUS RAICES

§ 20. Operaciones con los polinomios

La teoría de los determinantes y la teoría de los sistemas de ecuaciones lineales es un desarrollo directo de la rama del álgebra escolar que, comenzando por una ecuaci n de primer grado con una incógnita, nos lleva a los sistemas de dos y tres ecuaciones de primer grado con dos y tres incógnitas, respectivamente. Otra rama del álgebra elemental, considerada más importante, consiste en el *paso de una ecuación de primer grado con una incógnita a una ecuación cuadrática arbitraria, de nuevo con una incógnita, y después a unos tipos particulares de ecuaciones de tercero y cuarto grado.* Esta rama se desarrolla en una amplia y rica sección del álgebra superior dedicada al estudio de ecuaciones arbitrarias de n-ésimo grado con una incógnita. Esta sección del álgebra es la más antigua. El presente capítulo, así como algunos de los capítulos ulteriores del libro están relacionados con esta sección.

La forma general de una ecuación de n-ésimo grado (donde n es *cierto número entero positivo*) es

$$a_0x^n + a_1x^{n-1} + \ldots + a_{n-1}x + a_n = 0. \qquad (1)$$

Se supondrá que los coeficientes $a_0, a_1, \ldots, a_{n-1}, a_n$ son números complejos arbitrarios y que el *coeficiente superior* a_0 es diferente de cero.

Resolver la ecuación (1) significa hallar para la incógnita x *unos valores numéricos que satisfagan a esta ecuación, es decir,* que al sustituirlos en lugar de la incógnita, después de realizar todas las operaciones indicadas, reduzcan a cero el primer miembro de la ecuación (1).

Por otra parte, resulta conveniente sustituir el problema de la resolución de la ecuación (1) por el problema más general del estudio del primer miembro de esta ecuación

$$a_0x^n + a_1x^{n-1} + \ldots + a_{n-1}x + a_n, \qquad (2)$$

denominado *polinomio de grado n en la indeterminada x.* Hay que tener presente que ahora denominamos polinomio a una expresión

de la forma (2), o sea, a una suma de potencias no negativas de la indeterminada x, tomadas con ciertos coeficientes numéricos, y no a cualquier suma de monomios como ocurría en el álgebra elemental. En particular, no se denominarán polinomios las expresiones que contengan la indeterminada x con exponentes negativos o fraccionarios, por ejemplo, $2x^2 - \frac{1}{x} + 3$, o $ax^{-3} + bx^{-2} + cx^{-1} + d + ex + fx^2$, o bien, $x^{\frac{1}{2}} + 1$. Abreviadamente, los polinomios se designarán con las notaciones: $f(x)$, $g(x)$, $\varphi(x)$, etc.

Dos polinomios $f(x)$ y $g(x)$ se supondrán *iguales* (o *idénticamente iguales*), $f(x) = g(x)$, si son idénticos los coeficientes de potencias iguales de la indeterminada. En particular, un polinomio no puede ser idéntico a cero, si al menos uno de sus coeficientes es diferente de cero y, por lo tanto, el signo de igualdad que figura en la expresión de la ecuación de n-ésimo grado (1) no tiene que ver nada con la igualdad de los polinomios que acabamos de definir. El signo = que liga a los polinomios se debe entender como una identidad de los mismos.

Por consiguiente, el polinomio de n-ésimo grado (2) se debe interpretar como una expresión formal, completamente determinada por el conjunto de sus coeficientes $a_0, a_1, \ldots, a_n$, donde $a_0 \neq 0$. El significado exacto de estas palabras se aclarará mucho más tarde, en el cap. 10. Obsérvese que además de la expresión del polinomio en la forma (2), o sea, según las *potencias decrecientes de la indeterminada* x, se permitirán también otras expresiones obtenidas de (2) permutando los sumandos, como por ejemplo, la expresión según las *potencias crecientes de la indeterminada*.

Naturalmente, se podría interpretar el polinomio (2) desde el punto de vista del análisis matemático, o sea, como una función compleja de la variable compleja x. Sin embargo, se debe tener en cuenta que *dos funciones se suponen iguales cuando son iguales sus valores para cualesquiera valores de la variable* x. Está claro que dos polinomios que son iguales en el sentido algebraico formal indicado anteriormente, serán también iguales como funciones de x. Lo recíproco se demostrará en el § 24. Después de esto resultarán equivalentes los puntos de vista algebraico y teórico—funcional sobre el concepto de polinomio con coeficientes numé icos; por ahora tendremos que indicar cada vez el sentido que se da al concepto de polinomio. En el párrafo presente y en los dos que siguen se considerará el polinomio como una expresión algebraica formal.

Por supuesto, para cualquier número natural n existen polinomios de n grado. Examinando todos los polinomios posibles, además de los polinomios de primer grado, cuadráticos, cúbicos y etc., nos encontraremos con *polinomios de grado cero*, es decir, con núme-

ros complejos **diferentes de cero.** El número cero también se tomará como polinomio; éste es el único polinomio cuyo grado es indefinido.

A continuación definiremos las operaciones de adición y multiplicación para los polinomios de coeficientes complejos. Estas operaciones se introducirán del mismo modo que las operaciones con los polinomios de coeficientes reales, conocidas por el lector en el curso de álgebra elemental.

Dados los polinomios $f(x)$ y $g(x)$ de coeficientes complejos, expresados para mayor comodidad según las potencias crecientes de x:

$$f(x)=a_0+a_1x+\ldots+a_{n-1}x^{n-1}+a_nx^n, \qquad a_n\neq 0,$$
$$g(x)=b_0+b_1x+\ldots+b_{s-1}x^{s-1}+b_sx^s, \qquad b_s\neq 0,$$

donde $n\geqslant s$, se llamará *suma* al polinomio

$$f(x)+g(x)=c_0+c_1x+\ldots+c_{n-1}x^{n-1}+c_nx^n,$$

cuyos coeficientes se obtienen sumando los coeficientes respectivos de iguales potencias de la indeterminada en las expresiones de $f(x)$ y $g(x)$, o sea,

$$c_i=a_i+b_i, \quad i=0, 1, \ldots, n, \tag{3}$$

donde, para $n>s$, se tiene que suponer que los coeficientes b_{s+1}, b_{s+2}, ..., b_n son iguales a cero. El grado de la suma será igual a n, si n es mayor que s; pero, para $n=s$, puede ocurrir que éste sea menor que n, precisamente cuando $b_n=-a_n$.

Se llama *producto* de los polinomios $f(x)$ y $g(x)$ al polinomio

$$f(x)\cdot g(x)=d_0+d_1x+\ldots+d_{n+s-1}x^{n+s-1}+d_{n+s}x^{n+s},$$

cuyos coeficientes se determinan del modo siguiente:

$$d_i=\sum_{k+l=i} a_kb_l, \quad i=0, 1, \ldots, n+s-1, \; n+s, \tag{4}$$

o sea, el coeficiente d_i es el resultado de sumar todos los productos de aquellos coeficientes de los polinomios $f(x)$ y $g(x)$ la suma de cuyos índices es igual a i; en particular, $d_0=a_0b_0$, $d_1=a_0b_1+a_1b_0$, ..., $d_{n+s}=a_nb_s$. De la última igualdad resulta la desigualdad $d_{n+s}\neq 0$. Por consiguiente, el *grado del producto de dos polinomios es igual a la suma de sus grados.*

De aquí se deduce que *nunca será igual a cero el producto de polinomios, diferentes de cero.*

Para los polinomios, ¿qué propiedades poseen las operaciones introducidas? Las **propiedades conmutativa y asociativa de la suma** son consecuencia inmediata del cumplimiento de estas propiedades para la suma de los números, puesto que se suman los coeficientes

de cada potencia de la indeterminada por separado. La **resta** resulta posible: desempeña el papel del **cero** el número cero, que fue incluido como polinomio; el **opuesto** al polinomio $f(x)$, escrito anteriormente, es el polinomio

$$-f(x) = -a_0 - a_1x - \ldots - a_{n-1}x^{n-1} - a_nx^n.$$

La **propiedad conmutativa de la multiplicación** es consecuencia del cumplimiento de la propiedad conmutativa para el producto de los números y de que en la definición del producto de polinomios, los coeficientes de ambos factores $f(x)$ y $g(x)$ se empleen de un modo equivalente. La propiedad **asociativa** se demuestra del modo siguiente: si, además de los polinomios $f(x)$ y $g(x)$ escritos anteriormente, es dado también el polinomio

$$h(x) = c_0 + c_1x + \ldots + c_{t-1}x^{t-1} + c_tx^t, \quad c_t \neq 0,$$

el coeficiente de x^i, $i = 0, 1, \ldots, n+s+t$ en el producto $[f(x) \times \times g(x)]\, h(x)$ será el número

$$\sum_{j+m=i} \Big(\sum_{k+l=j} a_kb_l \Big) c_m = \sum_{k+l+m=i} a_kb_lc_m,$$

mientras que en el producto $f(x)\,[g(x)\,h(x)]$, será el número

$$\sum_{k+j=i} a_k \Big(\sum_{l+m=j} b_lc_m \Big) = \sum_{k+l+m=i} a_kb_lc_m.$$

que es igual a él.

Finalmente, el cumplimiento de la **ley distributiva** se deduce de la igualdad

$$\sum_{k+l=i} (a_k + b_k)\, c_l = \sum_{k+l=i} a_kc_l + \sum_{k+l=i} b_kc_l,$$

puesto que el primer miembro de ésta es el coeficiente de x^i en el polinomio $[f(x) + g(x)]\, h(x)$ y el segundo miembro es el coeficiente de la misma potencia de la indeterminada en el polinomio $f(x)\,h(x) + g(x)\,h(x)$.

Obsérvese que en el producto de los polinomios, el número 1, considerado como polinomio de grado cero, desempeña el papel de la **unidad**. Por otra parte, el *polinomio* $f(x)$ *posee un polinomio recíproco* $f^{-1}(x)$,

$$f(x)\, f^{-1}(x) = 1, \tag{5}$$

cuando, y sólo cuando, $f(x)$ *es un polinomio de grado cero*. En efecto, si $f(x)$ es un número a, diferente de cero, el polinomio recíproco es el número a^{-1}. Pero si $f(x)$ es de grado $n \geqslant 1$, el grado del primer miembro de la igualdad (5), en caso de que existiese el polinomio $f^{-1}(x)$ no sería menor de n, mientras que en el segundo miembro figura un polinomio de grado cero.

De aquí se deduce que *para el producto de polinomios no existe la operación inversa, la división*. En este sentido, el sistema de todos los polinomios de coeficientes complejos se parece al sistema de todos los números enteros. Esta analogía se manifiesta en que para los polinomios, al igual que para los números enteros, subsiste el **algoritmo de la división con resto***. Para el caso de los polinomios de coeficientes reales el lector ya conoce este algoritmo por el álgebra elemental. Pero, como consideramos ahora el caso de polinomios con *coeficientes complejos*, tendremos que hacer todos los enunciados due aquí se requieren y las demostraciones correspondientes.

Para cualesquiera dos polinomios $f(x)$ y $g(x)$ se pueden hallar unos polinomios $q(x)$ y $r(x)$, de tal manera que

$$f(x)=g(x)\,q(x)+r(x), \tag{6}$$

donde el grado de $r(x)$ es menor que el de $g(x)$, o bien, $r(x)=0$. Los polinomios $q(x)$ y $r(x)$ que satisfacen a esta condición se determinan unívocamente.

Demostremos primero la segunda parte del teorema. Supongamos que existen también unos polinomios $\overline{q}(x)$ y $\overline{r}(x)$ que satisfacen a la condición

$$f(x)=g(x)\,\overline{q}(x)+\overline{r}(x), \tag{7}$$

donde el grado de $\overline{r}(x)$ es de nuevo menor que el de $g(x)$**. Igualando entre sí los segundos miembros de las igualdades (6) y (7), se tiene:

$$g(x)\,[q(x)-\overline{q}(x)]=\overline{r}(x)-r(x).$$

El grado del segundo miembro de esta igualdad es menor que el de $g(x)$, mientras que si $q(x)-\overline{q}(x)\neq 0$, el grado del primer miembro sería mayor o igual al grado de $g(x)$. Por esto, tiene que ser $q(x)-\overline{q}(x)=0$, o sea, $q(x)=\overline{q}(x)$, de donde $r(x)=\overline{r}(x)$, como se quería demostrar.

Pasemos a demostrar la primera parte del teorema. Supongamos que n y s son los grados respectivos de los polinomios. Si $n<s$, se puede suponer que $q(x)=0$, $r(x)=f(x)$. Si $n\geqslant s$, aplicaremos el mismo método empleado en álgebra elemental para efectuar la división de polinomios con coeficientes reales, ordenados según las potencias decrecientes de la indeterminada. Sea

$$f(x)=a_0x^n+a_1x^{n-1}+\ldots+a_{n-1}x+a_n,\qquad a_0\neq 0,$$
$$g(x)=b_0x^s+b_1x^{s-1}+\ldots+b_{s-1}x+b_s,\qquad b_0\neq 0.$$

Poniendo

$$f(x)-\frac{a_0}{b_0}x^{n-s}g(x)=f_1(x), \tag{8}$$

* Denominada también división entera (*N del T.*).

** O bien $\overline{r}(x)=0$. En adelante, este caso no se va a excluir

se obtiene un polinomio cuyo grado es menor que n. Designemos este grado por n_1, y el coeficiente superior del polinomio $f_1(x)$, por a_{10}. Si todavía $n_1 \geqslant s$, ponemos

$$f_1(x) - \frac{a_{10}}{b_0} x^{n_1 - s} g(x) = f_2(x), \tag{8$_1$}$$

designamos con n_2 el grado y con a_{20}, el coeficiente superior del polinomio $f_2(x)$, poniendo después

$$f_2(x) - \frac{a_{20}}{b_0} x^{n_2 - s} g(x) = f_3(x), \tag{8$_2$}$$

y etc.

Como los grados de los polinomios $f_1(x)$, $f_2(x)$, ... decrecen, $n > n_1 > n_2 > \ldots$, después de repetir este proceso un número finito de veces, obtendremos un polinomio $f_k(x)$:

$$f_{k-1}(x) - \frac{a_{k-1,\,0}}{b_0} x^{n_{k-1} - s} g(x) = f_k(x), \tag{8$_{k-1}$}$$

cuyo grado n_k será menor que s, terminando así el proceso. Sumando ahora las igualdades (8), (8$_1$), ..., (8$_{k-1}$), se obtiene:

$$f(x) - \left(\frac{a_0}{b_0} x^{n-s} + \frac{a_{10}}{b_0} x^{n_1 - s} + \ldots + \frac{a_{k-1,\,0}}{b_0} x^{n_{k-1} - s}\right) g(x) = f_k(x),$$

o sea, que los polinomios

$$q(x) = \frac{a_0}{b_0} x^{n-s} + \frac{a_{10}}{b_0} x^{n_1 - s} + \ldots + \frac{a_{k-1,\,0}}{b_0} x^{n_{k-1} - s},$$
$$r(x) = f_k(x)$$

satisfacen verdaderamente a la igualdad (6), siendo realmente el grado de $r(x)$ menor que el de $g(x)$.

Obsérvese que el polinomio $q(x)$ se llama *cociente* de la división de $f(x)$ por $g(x)$ y $r(x)$, *resto* (o *residuo*) de esta división.

De la consideración del algoritmo de la división con resto, se establece fácilmente que si $f(x)$ y $g(x)$ *son polinomios de coeficientes reales, los coeficientes de todos los polinomios* $f_1(x)$, $f_2(x)$, ..., *y, por consiguiente, también los del cociente* $q(x)$ *y los del resto,* $r(x)$, *son reales.*

§ 21. Divisores. Máximo común divisor

Sean dados unos polinomios $f(x)$ y $\varphi(x)$, diferentes de cero, con coeficientes complejos. Si el resto de la división de $f(x)$ por $\varphi(x)$ es igual a cero, o como también se dice, si $f(x)$ *se divide* (o *es divisible*) por $\varphi(x)$, entonces el polinomio $\varphi(x)$ se llama *divisor* del polinomio $f(x)$.

El polinomio $\varphi(x)$ *es divisor del polinomio* $f(x)$ *si, y sólo si, existe un polinomio* $\psi(x)$ *que satisfaga a la igualdad*

$$f(x)=\varphi(x)\,\psi(x). \tag{1}$$

En efecto, si $\varphi(x)$ es divisor de $f(x)$, el cociente de la división de $f(x)$ por $\varphi(x)$ desempeña el papel de $\psi(x)$. Recíprocamente, supongamos que existe un polinomio $\psi(x)$ que satisface a la igualdad (1). De la unicidad de los polinomios $q(x)$ y $r(x)$ que satisfacen a la igualdad

$$f(x)=\varphi(x)\,q(x)+r(x).$$

demostrada en el párrafo anterior, y de la condición de que el grado de $r(x)$ es menor que el de $\varphi(x)$, se deduce que en este caso el cociente de la división de $f(x)$ por $\varphi(x)$ es igual a $\psi(x)$ y el resto es igual a cero.

Se comprende que, cumpliéndose la igualdad (1), $\psi(x)$ es también divisor de $f(x)$. Luego, es evidente que el grado de $\varphi(x)$ no es superior al de $f(x)$.

Obsérvese que, si el polinomio $f(x)$ y su divisor $\varphi(x)$ tienen ambos coeficientes racionales o reales, el polinomio $\psi(x)$ también tiene los coeficientes racionales o, respectivamente, reales, puesto que éste se halla mediante el algoritmo de la división. Por supuesto, un polinomio de coeficientes racionales o reales puede poseer también divisores cuyos coeficientes no sean todos racionales o, respectivamente, reales. Esto se observa, por ejemplo, en la igualdad

$$x^2+1=(x-i)(x+i).$$

Señalemos algunas propiedades fundamentales de la divisibilidad de los polinomios que tendrán numerosas aplicaciones a continuación.

I. *Si* $f(x)$ *es divisible por* $g(x)$ y $g(x)$ es divisible por $h(x)$, *entonces* $f(x)$ *es divisible por* $h(x)$.

En efecto, por la condición $f(x)=g(x)\,\varphi(x)$ y $g(x)=$ $=h(x)\,\psi(x)$, y, por lo tanto, $f(x)=h(x)\,[\psi(x)\,\varphi(x)]$.

II. *Si* $f(x)$ *y* $g(x)$ *es divisible por* $\varphi(x)$, *su suma y diferencia también es divisible por* $\varphi(x)$.

En efecto, de las igualdades $f(x)=\varphi(x)\,\psi(x)$ y $g(x)=$ $=\varphi(x)\,\chi(x)$ resulta: $f(x)\pm g(x)=\varphi(x)\,[\psi(x)\pm\chi(x)]$.

III. *Si* $f(x)$ *es divisible por* $\varphi(x)$, *el producto de* $f(x)$ *por cualquier po inomio* $g(x)$ *también es divisible por* $\varphi(x)$.

En efecto, si $f(x)=\varphi(x)\,\psi(x)$, se tiene: $f(x)\,g(x)=$ $=\varphi(x)\,[\psi(x)\,g(x)]$.

De II y III se deduce la siguiente propiedad:

IV. *Si cada uno de los polinomios* $f_1(x), f_2(x), \ldots, f_k(x)$ *es divisible por* $\varphi(x)$, *el polinomio*

$$f_1(x)\,g_1(x)+f_2(x)\,g_2(x)+\ldots+f_k(x)\,g_k(x),$$

donde $g_1(x)$, $g_2(x)$, ... $g_k(x)$ *son unos polinomios arbitrarios, también es divisible por* $\varphi(x)$.

V. *Todo polinomio* $f(x)$ *es divisible por cualquier polinomio de grado cero.*

En efecto, si $(x) = a_0x^n + a_1x^{n-1} + \ldots + a_n$, y c es un número arbitrario, diferente de cero, o sea, un polinomio arbitrario de grado cero, entonces

$$f(x) = c\left(\frac{a_0}{c}x^n + \frac{a_1}{c}x^{n-1} + \ldots + \frac{a_n}{c}\right).$$

VI. *Si* $f(x)$ *es divisible por* $\varphi(x)$, $f(x)$ *es también divisible por* $c\varphi(x)$, *donde* c *es un número arbitrario, diferente de cero.*

En efecto, de la igualdad $f(x) = \varphi(x)\psi(x)$ resulta la igualdad $f(x) = [c\varphi(x)]\cdot[c^{-1}\psi(x)]$.

VII. *Los polinomios* $cf(x)$, $c \neq 0$, *y sólo éstos, son los divisores del polinomio* $f(x)$ *que tienen el mismo grado que* $f(x)$.

En efecto, $f(x) = c^{-1}[cf(x)]$, o ea, $f(x)$ es divisible por $cf(x)$.

Si, por otra parte, $f(x)$ es divisible por $\varphi(x)$, coincidiendo los grados de $f(x)$ y $\varphi(x)$, el grado del cociente de la división de $f(x)$ por $\varphi(x)$ tiene que ser igual a cero, es decir, $f(x) = d\varphi(x)$, $d \neq 0$, de donde $\varphi(x) = d^{-1}f(x)$.

De aquí resulta la siguiente propiedad:

VIII. *Los polinomios* $f(x)$ *y* $g(x)$ *son divisibles entre sí cuando, y sólo cuando,* $g(x) = cf(x)$, $c \neq 0$.

Finalmente, de VIII y I resulta la propiedad:

IX. *Todo divisor de uno de los dos polinomios* $f(x)$ y $cf(x)$, donde $c \neq 0$, es divisor del otro.

Máximo común divisor. Sean dados unos polinomios arbitrarios, $f(x)$ y $g(x)$. El polinomio $\varphi(x)$ se llama *divisor común* de $f(x)$ y $g(x)$, si es divisor de cada uno de estos polinomios. La propiedad *V* (véase más arriba) muestra que todos los polinomios de grado cero pertenecen al conjunto de los divisores comunes de los polinomios $f(x)$ y $g(x)$. Si éstos no tienen más divisores comunes, se dice que *son primos entre sí*.

En el caso general, los polinomios $f(x)$ y $g(x)$ pueden poseer divisores dependientes de x. Introduzcamos ahora el concepto de *máximo* común divisor de estos polinomios.

Sería incómodo tomar tal definición, según la cual el máximo común divisor de los polinomios sería el común divisor de mayor grado. Por una parte, no sabemos todavía si $f(x)$ y $g(x)$ pueden poseer o no muchos divisores comunes distintos de mayor grado, que no sólo se diferencien entre sí en un factor de grado cero, por lo que esta definición resultaría muy indeterminada. Por otra parte, el lector ya conocerá por la aritmética elemental la forma de obtener el máximo común divisor de números enteros, y sabrá que el máximo común divisor 6 de los números enteros 12 y 18, no sólo es el

mayor de los divisores comunes de estos números, sino que también es divisible por cualquier otro de sus divisores comunes; en efecto, los demás divisores comunes de los números 12 y 18 son: 1, 2, 3, —1, —2, —3, —6.

Por consiguiente, para el caso de los polinomios, aceptaremos la siguiente definición:

Se llama *máximo común divisor* de los polinomios $f(x)$ y $g(x)$, diferentes de cero, al polinomio $d(x)$ que es común divisor y que a la vez es divisible por cualquier otro divisor común de estos polinomios. El *máximo común* divisor de los polinomios $f(x)$ y $g(x)$ se designa con la notación $(f(x), g(x))$.

Con esta definición no se aclara el problema de la existencia del máximo común divisor para cualesquiera polinomios $f(x)$ y $g(x)$. Ahora se dará una respuesta positiva a este problema. A la vez, se señalará un método para hallar el máximo común divisor de los polinomios dados. Por supuesto, aquí no se puede emplear el método con el que ordinariamente se halla el máximo común divisor de los números enteros, puesto que para los polinomios no tenemos por ahora nada parecido a la descomposición del número entero en un producto de factores primos. Sin embargo, para los números enteros existe también otro método, denominado *algoritmo de las divisiones sucesivas* o *algoritmo de Euclides*; este método también puede aplicarse a los polinomios.

El algoritmo de Euclides para los polinomios consiste en lo siguiente. Sean dados los polinomios $f(x)$ y $g(x)$. Se divide $f(x)$ por $g(x)$ y se obtiene por lo general un resto $r_1(x)$. Se divide luego $g(x)$ por $r_1(x)$ y se obtiene el resto $r_2(x)$, se divide $r_1(x)$ por $r_2(x)$, etc. Como los grados de los restos van disminuyendo todo el tiempo, en esta cadena de divisiones sucesivas llegaremos a un lugar en el que la división será exacta y el proceso se terminará. *El resto $r_k(x)$, por el que se divide exactamente el resto anterior $r_{k-1}(x)$, será el máximo común divisor de los polinomios $f(x)$ y $g(x)$.*

Para la demostración, escribamos lo expuesto en las líneas anteriores en forma de la siguiente cadena de igualdades:

$$\left.\begin{aligned} f(x) &= g(x)\,q_1(x) + r_1(x), \\ g(x) &= r_1(x)\,q_2(x) + r_2(x), \\ r_1(x) &= r_2(x)\,q_3(x) + r_3(x), \\ &\dots\dots\dots\dots\dots\dots \\ r_{k-3}(x) &= r_{k-2}(x)\,q_{k-1}(x) + r_{k-1}(x), \\ r_{k-2}(x) &= r_{k-1}(x)\,q_k(x) + r_k(x), \\ r_{k-1}(x) &= r_k(x)\,q_{k+1}(x). \end{aligned}\right\} \qquad (2)$$

La última igualdad muestra que $r_k(x)$ es divisor de $r_{k-1}(x)$. De aquí resulta que ambos sumandos del segundo miembro de la

penúltima igualdad son divisibles por $r_k(x)$, y por lo tanto, $r_k(x)$ también es divisor de $r_{k-2}(x)$. A continuación, subiendo del mismo modo a los otros renglones, obtenemos que $r_k(x)$ también es divisor de $r_{k-3}(x)$, ..., $r_2(x)$, $r_1(x)$. De aquí, en virtud de la segunda igualdad, resulta que $r_k(x)$ es divisor de $g(x)$, de donde, en virtud de la primera igualdad, también es divisor de $f(x)$. Por lo tanto, $r_k(x)$ es un divisor común de $f(x)$ y $g(x)$.

Consideremos ahora un divisor común arbitrario $\varphi(x)$ de los polinomios $f(x)$ y $g(x)$. Como el primer miembro y el primer sumando del segundo miembro de la primera de las igualdades (2) son divisibles por $\varphi(x)$, $r_1(x)$ también es divisible por $\varphi(x)$. Pasando a la segunda y a las siguientes igualdades, obtenemos del mismo modo que todos los polinomios $r_2(x)$, $r_3(x)$, ..., son divisibles por $\varphi(x)$. Si, finalmente, se ha demostrado que $r_{k-2}(x)$ y $r_{k-1}(x)$ son divisibles por $\varphi(x)$, de la penúltima igualdad obtenemos que $r_k(x)$ es divisible por $\varphi(x)$. Por lo tanto, $r_k(x)$ es verdaderamente el máximo común divisor de $f(x)$ y $g(x)$.

Por consiguiente, hemos demostrado que dos polinomios cualesquiera poseen máximo común divisor y hemos obtenido un método para su cálculo. Este método muestra que, *si los polinomios $f(x)$ y $g(x)$ tienen ambos coeficientes racionales o reales, los coeficientes de su máximo común divis r también son racionales o, respectivamente, reales*, a pesar de que estos polinomios pueden tener divisores cuyos coeficientes no sean todos racionales (reales). Así, pues, los polinomios con coeficientes racionales

$$f(x)=x^3-3x^2-2x+6, \quad g(x)=x^3+x^2-2x-2$$

tienen un máximo común divisor x^2-2 con coeficientes racionales, a pesar de que tienen un común divisor $x-\sqrt{2}$ cuyos coeficientes no son todos racionales.

Si $d(x)$ es el máximo común divisor de los polinomios $f(x)$ y $g(x)$, como muestran las propiedades VIII y IX (véase la pág. 139), por máximo común divisor se podría tomar también el polinomio $cd(x)$, donde c es un número arbitrario, diferente de cero. En otras palabras, el *máximo común divisor de dos polinomios se determina salvo un factor de grado cero*. En virtud de esto, se puede convenir en que el **coeficiente superior del máximo común divisor de dos polinomios sea siempre igual a la unidad.** Aplicando esta condición, se puede afirmar que *dos polinomios son primos entre sí cuando, y sólo cuando, su máximo común divisor es igual a la unidad.* En efecto, por máximo común divisor de dos polinomios, primos entre sí, se puede tomar cualquier número diferente de cero; pero si este número se multiplica por su elemento recíproco, obtenemos la unidad.

Ejemplo. Hallar el máximo común divisor de los polinomios

$$f(x)=x^4+3x^3-x^2-4x-3,\quad g(x)=3x^3+10x^2+2x-3.$$

Para evitar coeficientes fraccionarios, al aplicar el algoritmo de Euclides a los polinomios con coeficientes enteros, se puede multiplicar el dividendo o simplificar el divisor por cualquier número diferente de cero, no sólo al comenzar alguna de las divisiones sucesivas, sino también durante el proceso de la división misma. Naturalmente, esto conducirá a una alteración del cociente, pero los restos que nos interesan adquirirán solamente un factor de grado cero, lo que, como ya sabemos, al buscar el máximo común divisor, es admisible.

Dividimos $f(x)$ por $g(x)$, multiplicando previamente $f(x)$ por 3:

$$\begin{array}{r|l} 3x^4+9x^3-3x^2-12x-9 & 3x^3+10x^2+2x-3 \\ \underline{3x^4+10x^3+2x^2-3x} & x+1 \\ -x^3-5x^2-9x-9 & \end{array}$$

(multiplicamos por -3)

$$\begin{array}{r} 3x^3+15x^2+27x+27 \\ \underline{3x^3+10x^2+2x-3} \\ 5x^2+25x+30. \end{array}$$

Por lo tanto, después de simplificar por 5, el primer resto es $r_1(x)=$ $=x^2+5x+6$. Dividimos el polinomio $g(x)$ por éste:

$$\begin{array}{r|l} 3x^3+10x^2+2x-3 & x^2+5x+6 \\ \underline{3x^3+15x^2+18x} & 3x-5 \\ -5x^2-16x-3 & \\ \underline{-5x^2-25x-30} & \\ 9x+27. & \end{array}$$

Por consiguiente, después de simplificar por 9, el segundo resto es: $r_2(x)=x+3$. Como

$$r_1(x)=r_2(x)\,(x+2),$$

$r_2(x)$ será el último resto, por el que se divide exactamente el resto anterior. Por lo tanto, éste es el máximo común divisor:

$$(f(x),\ g(x))=x+3.$$

Apliquemos el algoritmo de Euclides para la demostración del **teorema** siguiente:

Si $d(x)$ *es el máximo común divisor de los polinomios* $f(x)$ *y* $g(x)$, *existen tales polinomios* $u(x)$ *y* $v(x)$ *que*

$$f(x)\,u(x)+g(x)\,v(x)=d(x). \qquad (3)$$

Siempre se puede suponer que, si los grados de los polinomios $f(x)$ *y* $g(x)$ *son mayores que cero, el grado de* $u(x)$ *es menor que el grado de* $g(x)$ *y el grado de* $v(x)$ *es menor que el grado de* $f(x)$.

La demostración está basada en las igualdades (2). Si se tiene en cuenta que $r_k(x)=d(x)$ y se pone $u_1(x)=1$, $v_1(x)=-q_k(x)$, según la penúltima de las igualdades (2), se tiene:

$$d(x)=r_{k-2}(x)\,u_1(x)+r_{k-1}(x)\,v_1(x).$$

Poniendo aquí la expresión de $r_{k-1}(x)$ mediante $r_{k-3}(x)$ y $r_{k-2}(x)$, se obtiene de la igualdad anterior (2):

$$d(x)=r_{k-3}(x)\,u_2(x)+r_{k-2}(x)\,v_2(x),$$

donde, evidentemente, $u_2(x)=v_1(x)$, $v_2(x)=u_1(x)-v_1(x)\,q_{k-1}(x)$. Continuando el ascenso por las igualdades (2), se llegará, finalmente, a la igualdad (3) que se quería demostrar.

Para la demostración de la segunda afirmación del teorema, supongamos que se han hallado ya los polinomios $u(x)$ y $v(x)$ que satisfacen a la gualdad (3), pero que el grado de $u(x)$ es, por ejemplo, mayor o igual al grado de $g(x)$. Dividamos $u(x)$ por $g(x)$:

$$u(x)=g(x)\,q(x)+r(x),$$

donde el grado de $r(x)$ es menor que el grado de $g(x)$, e introduzcamos esta expresión en (3). Se obtiene la igualdad.

$$f(x)\,r(x)+g(x)\,[v(x)+f(x)]q(x)]=d(x).$$

El grado del factor que figura con $f(x)$ es ya menor que el grado de $g(x)$. Por otra parte, el grado del polinomio que figura entre corchetes es menor que el grado de $f(x)$, puesto que, en caso contrario, el grado del segundo sumando del primer miembro no sería menor que el grado del producto $g(x)\,f(x)$, y como el grado del primer sumando es menor que el grado de este producto, todo el primer miembro sería de grado mayor o igual a $g(x)\,f(x)$, mientras que según nuestra suposición, el polinomio $d(x)$ es de menor grado.

Así el teorema queda demostrado. A la vez, tenemos que, si los polinomios $f(x)$ y $g(x)$ tienen coeficientes racionales o reales, los polinomios $u(x)$ y $v(x)$ que satisfacen a la igualdad (3) se pueden elegir de modo que sus coeficientes sean racionales o, respectivamente, reales.

Ejemplo. Hallemos los polinomios $u(x)$ y $v(x)$ que satisfacen a la igualdad (3), si

$$f(x)=x^3-x^2+3x-10,\quad g(x)=x^3+6x^2-9x-14.$$

Apliquemos el algoritmo de Euclides a estos polinomios: ahora, al efectuar las divisiones ya no se puede permitir ninguna alteración de los cocientes, puesto que éstos se emplean para hallar los polinomios $u(x)$ y $v(x)$. Obtenemos el siguiente sistema de igualdades:

$$f(x)=g(x)+(-7x^2+12x+4);$$
$$g(x)=(-7x^2+12x+4)\left(-\frac{1}{7}x-\frac{54}{49}\right)+\frac{235}{49}(x-2);$$
$$-7x^2+12x+4=(x-2)(-7x-2).$$

De aquí sale que $(f(x),\,g(x))=x-2$ y que

$$u(x)=\frac{7}{235}x+\frac{54}{235},\quad v(x)=-\frac{7}{235}x-\frac{5}{235}.$$

Aplicando ahora el teorema demostrado a polinomios, primos entre sí, obtenemos el siguiente resultado:

Los polinomios $f(x)$ y $g(x)$ *son primos entre sí cuando, y sólo cuando, existen unos polinomios* $u(x)$ y $v(x)$ *que satisfacen a la igualdad*

$$f(x)\,u(x)+g(x)\,v(x)=1. \tag{4}$$

Basándose en este resultado se pueden demostrar unos cuantos **teoremas sobre los polinomios primos entre sí**, que, aunque sencillos, son importantes:

a) *Si el polinomio* $f(x)$ *es primo con cada uno de los polinomios* $\varphi(x)$ y $\psi(x)$, *también es primo con su producto.*

En efecto, según (4), existen unos polinomios $u(x)$ y $v(x)$ tales que

$$f(x)\,u(x)+\varphi(x)\,v(x)=1.$$

Multiplicando esta igualdad por $\psi(x)$, obtenemos:

$$f(x)\,[u(x)\,\psi(x)]+[\varphi(x)\,\psi(x)]\,v(x)=\psi(x),$$

de donde se deduce que todo común divisor de $f(x)$ y $\varphi(x)\,\psi(x)$ es también divisor de $\psi(x)$; sin embargo, según la condición, $(f(x), \psi(x))=1$.

b) *Si el producto de los polinomios* $f(x)$ *y* $g(x)$ *es divisible por* $\varphi(x)$, *pero* $f(x)$ *y* $\varphi(x)$ *son primos entre sí*, $g(x)$ *es divisible por* $\varphi(x)$.

En efecto, multiplicando la igualdad

$$f(x)\,u(x)+\varphi(x)\,v(x)=1$$

por $g(x)$, obtenemos:

$$[f(x)\,g(x)]\,u(x)+\varphi(x)\,[v(x)\,g(x)]=g(x).$$

Ambos sumandos del primer miembro de esta igualdad son divisibles por $\varphi(x)$; por consiguiente, también $g(x)$ es divisible por $\varphi(x)$.

c) *Si el polinomio* $f(x)$ *es divisible por cada uno de los polinomios* $\varphi(x)$ *y* $\psi(x)$, *que son primos entre sí, entonces* $f(x)$ *también es divisible por su producto.*

En efecto, $f(x)=\varphi(x)\,\overline{\varphi}(x)$, o sea, el producto que figura en el segundo miembro es divisible por $\psi(x)$. Por esto, según b), $\overline{\varphi}(x)$ es divisible por $\psi(x)$, $\overline{\varphi}(x)=\psi(x)\,\overline{\psi}(x)$, de donde $f(x)=$ $=[\varphi(x)\,\psi(x)]\,\overline{\psi}(x)$.

La definición de máximo común divisor se puede generalizar al caso de cualquier sistema finito de polinomios. Se llama *máximo común divisor* de los polinomios $f_1(x), f_2(x), \ldots, f_s(x)$ a su divisor común que es divisible por cualquier otro divisor común de los mismos. La existencia del máximo común divisor para cualquier

sistema finito de polinomios es consecuencia del siguiente **teorema,** que facilita también un procedimiento para su cálculo.

El máximo común divisor de los polinomios $f_1(x), f_2(x), \ldots, f_s(x)$ *es igual al máximo común divisor del polinomio* $f_s(x)$ *y del máximo común divisor de los polinomios* $f_1(x), f_2(x), \ldots, f_{s-1}(x)$.

En efecto, el teorema es evidente para $s = 2$. Por esto, supondremos que el teorema subsiste para el caso $s - 1$, o sea, que, en particular, ya está demostrada la existencia del máximo común divisor $d(x)$ de los polinomios $f_1(x), f_2(x), \ldots, f_{s-1}(x)$. Designemos mediante $\overline{d}(x)$ el máximo común divisor de los polinomios $d(x)$ y $f_s(x)$. Es evidente que éste es un divisor común de todos los polinomios dados. Por otra parte, cualquier otro divisor común de estos polinomios es también divisor de $d(x)$ y, por lo tanto, de $\overline{d}(x)$.

En particular, se dice que $f_1(x), f_2(x), \ldots, f_s(x)$ es un sistema de polinomios, *primos entre sí*, si los únicos divisores comunes de ellos son los polinomios de grado cero, o sea, si su máximo común divisor es igual a 1. Si $s > 2$, puede ocurrir que estos polinomios no sean primos entre sí dos a dos. Así, pues, los polinomios

$$f(x) = x^3 - 7x^2 + 7x + 15,\ \ g(x) = x^2 - x - 20,$$
$$h(x) = x^3 + x^2 - 12x$$

son primos entre sí, a pesar de que

$$(f(x),\ g(x)) = x - 5,\ \ (f(x),\ h(x)) = x - 3,\ \ (g(x),\ h(x)) = x + 4.$$

El lector obtendrá fácilmente la generalización de los teoremas a) — c), demostrados anteriormente, sobre los polinomios primos entre sí, para el caso de cualquier número finito de polinomios

§ 22. Las raíces de los polinomios

En el § 20 nos encontramos con los valores de un polinomio, cuando se hablaba del punto de vista teórico-funcional del concepto de polinomio. Recordemos la definición.

Si

$$f(x) = a_0x^n + a_1x^{n-1} + \ldots + a_n, \tag{1}$$

es un polinomio y c es un número, el número

$$f(c) = a_0c^n + a_1c^{n-1} + \ldots + a_n,$$

obtenido por la sustitución de la indeterminada x por el número c, en la expresión (1) de $f(x)$, y por la realización consiguiente de las operaciones indicadas, se denomina *valor del polinomio* $f(x)$ *para* $x = c$. Se comprende que, si $f(x) = g(x)$, en el sentido de la igualdad algebraica de polinomios definida en el § 20, entonces $f(c) = g(c)$ para cualquier c.

Fácilmente se ve también que si

$$\varphi(x) = f(x) + g(x), \quad \psi(x) = f(x)\,g(x),$$

se tiene

$$\varphi(c) = f(c) + g(c), \quad \psi(c) = f(c)\,g(c).$$

En otras palabras, considerando a los polinomios desde el punto de vista teórico-funcional, la suma y el producto de los polinomios definidas en el § 20, se convierten en la suma y el producto de funciones, consideradas en el sentido de la suma y producto de los valores respectivos de estas funciones.

Si $f(c) = 0$, o sea, si el polinomio $f(x)$ se anula al *sustituir el número c en lugar de la indeterminada, c se llama raíz* del polinomio $f(x)$* (o de la ecuación $f(x) = 0$). Ahora se demostrará que este concepto está relacionado con la teoría de la divisibilidad de los polinomios, estudiada en el párrafo anterior.

Si se divide el polinomio $f(x)$ por un polinomio arbitrario de primer grado (o como se dirá a continuación, por un *polinomio lineal*), el resto será un polinomio de grado cero o bien será igual a cero, es decir, siempre será un número r. Aplicando el **teorema** que sigue es fácil hallar este resto sin realizar la división (se supone que el polinomio lineal es de la forma $x - c$).

El resto de la división de un polinomio $f(x)$ por un polinomio lineal $x - c$ es igual al valor $f(c)$ que toma el polinomio $f(x)$ para $x = c$.

En efecto, sea

$$f(x) = (x - c)\,q(x) + r.$$

Tomando los valores de ambos miembros de esta igualdad para $x = c$, obtenemos:

$$f(c) = (c - c)\,q(c) + r = r,$$

lo cual demuestra el teorema.

De aquí se deduce la importante **conclusión**:

El número c es raíz del polinomio $f(x)$ cuando, y sólo cuando, $f(x)$ es divisible por $x - c$.

Por otra parte, es evidente que si $f(x)$ es divisible por algún polinomio de primer grado $ax + b$, es divisible también por el polinomio $x - \left(-\frac{b}{a}\right)$, o sea, por un polinomio de la forma $x - c$. De este modo, la *averiguación de las raíces del polinomio $f(x)$ es equivalente a la averiguación de sus divisores lineales.*

En virtud de lo expuesto anteriormente, el siguiente método de división de un polinomio $f(x)$ por el binomio lineal $x - c$ es de especial interés, pues es más simple que el algoritmo general de di-

* También se dice que c es un cero del polinomio $f(x)$. (*Nota del T.*)

visión de los polinomios. Este método se denomina **regla de Horner.** Sea

$$f(x) = a_0x^n + a_1x^{n-1} + a_2x^{n-2} + \ldots + a_n \qquad (2)$$

y supongamos que

$$f(x) = (x - c)\, q(x) + r, \qquad (3)$$

donde

$$q(x) = b_0x^{n-1} + b_1x^{n-2} + b_2x^{n-3} + \ldots + b_{n-1}.$$

Igualando en (3) los coeficientes de potencias iguales de x, obtenemos:

$$\begin{aligned} a_0 &= b_0, \\ a_1 &= b_1 - cb_0, \\ a_2 &= b_2 - cb_1, \\ &\ldots\ldots\ldots\ldots \\ a_{n-1} &= b_{n-1} - cb_{n-2}, \\ a_n &= r - cb_{n-1}. \end{aligned}$$

De aquí se deduce que $b_0 = a_0$, $b_k = cb_{k-1} + a_k$, $k = 1, 2, \ldots, n-1$, o sea, se obtiene **el coeficiente b_k multiplicando el coeficiente anterior b_{k-1} por c y agregándole el coeficiente correspondiente a_k**; finalmente, $r = cb_{n-1} + a_n$, es decir, el resto r, que como ya sabemos es igual a $f(c)$, se obtiene por la misma regla. Por lo tanto, los coeficientes del cociente y el resto se pueden obtener sucesivamente mediante unos cálculos del mismo tipo; éstos se realizan de acuerdo a un esquema, como se muestra en los siguientes ejemplos:

1. Dividir $f(x) = 2x^5 - x^4 - 3x^3 + x - 3$ por $x - 3$.

Formemos una tabla colocando sobre la raya los coeficientes del polinomio $f(x)$; bajo la raya se colocan los coeficientes correspondientes del cociente y del resto que se calculan sucesivamente y, a la izquierda, a un lado, el valor dado de c:

$$3 \left| \begin{array}{llllll} 2 & -1 & -3 & 0 & 1 & -3 \\ \hline 2, & 3\cdot 2-1=5, & 3\cdot 5-3=12, & 3\cdot 12+0=36, & 3\cdot 36+1=109, & 3\cdot 109-3=324 \end{array} \right. .$$

Por lo tanto, el coeficiente buscado es

$$q(x) = 2x^4 + 5x^3 + 12x^2 + 36x + 109,$$

y el resto, $r = f(3) = 324$.

2. Dividir $f(x) = x^4 - 8x^3 + x^2 + 4x - 9$ por $x + 1$.

$$-1 \left| \begin{array}{rrrrr} 1 & -8 & 1 & 4 & -9 \\ \hline 1 & -9 & 10 & -6 & -3. \end{array} \right.$$

Por consiguiente, el cociente es

$$q(x) = x^3 - 9x^2 + 10x - 6,$$

y el resto ($r = f(-1)$) $= -3$

Estos ejemplos muestran que la *regla de Horner se puede utilizar también para calcular rápidamente el valor del polinomio para un valor dado de la indeterminada.*

Raíces múltiples. Si c es una raíz del polinomio $f(x)$, o sea, si $f(c) = 0$, entonces, como ya sabemos $f(x)$ es divisible por $x - c$. Puede ocurrir que $f(x)$ no sólo sea divisible por la primera potencia del binomio lineal $x - c$, sino también por potencias superiores. De todos modos, siempre existirá un número natural k tal que $f(x)$ sea divisible por $(x - c)^k$, pero no por $(x - c)^{k+1}$. En consecuencia,

$$f(x) = (x - c)^k \varphi(x),$$

en donde el polinomio $\varphi(x)$ ya no es divisible por $x - c$, o sea, el número c no es raíz de $\varphi(x)$. El número k se llama orden de *multiplicidad* de la raíz c del polinomio $f(x)$ y el número c, *raíz múltiple* de este polinomio de orden k. Si $k = 1$, se dice que c es una raíz *simple*.

El concepto de raíz múltiple está estrechamente ligado con el concepto de derivada del polinomio. Como estamos estudiando los polinomios con coeficientes complejos cualesquiera, no podemos utilizar directamente el concepto de derivada que se introdujo en el curso de análisis matemático. Todo lo que se diga a continuación se debe considerar como una **definición** de la derivada de un polinomio, independiente del curso de análisis.

Sea dado un polinomio de n-ésimo grado

$$f(x) = a_0x^n + a_1x^{n-1} + \ldots + a_{n-1}x + a_n$$

con cualesquiera coeficientes complejos. El polinomio de $(n-1)$-ésimo grado,

$$f'(x) = na_0x^{n-1} + (n-1)a_1x^{n-2} + \ldots + 2a_{n-2}x + a_{n-1},$$

se llama *derivada* (o *derivada primera*) del polinomio $f(x)$. La derivada de cero y de un polinomio de grado cero se supone igual a cero. La derivada de la derivada primera se llama *derivada segunda* del polinomio $f(x)$ y se designa con $f''(x)$, etc. Es evidente que

$$f^{(n)}(x) = n!\ a_0,$$

por lo cual $f^{(n+1)}(x) = 0$, es decir, la *$(n+1)$-ésima derivada de un polinomio de n-ésimo grado es igual a cero.*

En el caso de polinomios de coeficientes complejos no podemos utilizar las propiedades de la derivada que fueron demostradas en el curso de análisis para los polinomios de coeficientes reales, sino que tenemos que demostrar de nuevo estas propiedades utilizando solamente la definición de derivada dada anteriormente. Aquí nos interesan las siguientes propiedades que, como se suele decir, re-

presentan las fórmulas del derivación de la suma y del producto:

$$(f(x)+g(x))' = f'(x)+g'(x), \tag{4}$$

$$(f(x)\cdot g(x))' = f(x)g'(x)+f'(x)g(x). \tag{5}$$

Estas fórmulas se comprueban fácilmente mediante un cálculo directo, tomando por $f(x)$ y $g(x)$ dos polinomios arbitrarios y aplicando la definición de derivada dada anteriormente; recomendamos al lector hacerlo.

La fórmula (5) se generaliza sin dificultad al caso de un producto de cualquier número finito de factores, de donde, de un modo ordinario, se puede deducir la fórmula para la derivada de una potencia:

$$(f^k(x))' = kf^{k-1}(x)f'(x). \tag{6}$$

Nuestro propósito es demostrar el teorema siguiente:

Si el número c es una raíz k-múltiple del polinomio $f(x)$, entonces, para $k>1$, éste será una raíz $(k-1)$-múltiple de la derivada primera de este polinomio; si $k=1$, el número c no será raíz de $f'(x)$.

En efecto, sea

$$f(x) = (x-c)^k\varphi(x), \quad k \geqslant 1, \tag{7}$$

donde $\varphi(x)$ ya no es divisible por $x-c$. Derivando la igualdad (7), se obtiene:

$$f'(x) = (x-c)^k\varphi'(x) + k(x-c)^{k-1}\varphi(x) =$$
$$= (x-c)^{k-1}[(x-c)\varphi'(x)+k\varphi(x)].$$

El primer término de la suma que figura entre los corchetes es divisible por $x-c$, mientras que el segundo no es divisible por éste; por consiguiente, toda esta suma no puede ser divisible por $x-c$. Teniendo en cuenta que el cociente de la división de $f(x)$ por $(x-c)^{k-1}$ se determina unívocamente, resulta que $(x-c)^{k-1}$ es la máxima potencia del binomio $x-c$ por la cual es divisible el polinomio $f'(x)$, como se quería demostrar.

Aplicando unas cuantas veces este teorema, obtenemos que la *raíz k-múltiple del polinomio $f(x)$ es raíz $(k-s)$-múltiple de la s-ésima derivada de este polinomio $(k \geqslant s)$ y que por primera vez no será raíz para la k-ésima derivada de $f(x)$.*

§ 23. Teorema fundamental

Al estudiar en el párrafo anterior las raíces de los polinomios, no planteamos el problema de la existencia de raíces para cualquier polinomio. Se sabe que existen polinomios de coeficientes reales que no tienen raíces reales como, por ejemplo: x^2+1. Se podría esperar que existiesen polinomios que no tuviesen raíces incluso

entre los números complejos, sobre todo si se consideran polinomios con cualesquiera coeficientes complejos. Si esto fuese así, se necesitaría una ampliación ulterior del sistema de los números complejos. Sin embargo, en realidad, subsiste el siguiente **teorema fundamental del álgebra de los números complejos:**

Todo polinomio de cualesquiera coeficientes numéricos, cuyo grado no sea menor que la unidad, tiene por lo menos una raíz, generalmente, compleja.

Este teorema es uno de los adelantos más grandiosos de toda la matemática y encuentra aplicación en las más diversas ramas de la ciencia. En particular, en él se basa toda la teoría ulterior de los polinomios con coeficientes numéricos. Por esta razón, le llamaban antes (y a veces ahora también le llaman) «teorema fundamental del álgebra superior». No obstante, el teorema fundamental no es puramente algebraico. Todas sus demostraciones (después de Gauss, que fue el primero en demostrar este teorema a fines del siglo XVIII, se hallaron muchas otras demostraciones), en tal o cual grado, emplean las llamadas propiedades topológicas de los números reales y complejos, o sea, las propiedades que están ligadas a la continuidad.

En la demostración que se va a exponer ahora, el polinomio $f(x)$ de coeficientes complejos se va a considerar como una función de la variable compleja x. Por lo tanto, x puede tomar cualesquiera valores complejos, o como suele decirse, teniendo en cuenta el método de construcción de los números complejos expuesto en el § 17, la variable x varía en el *plano complejo*. Los valores de la función $f(x)$ también serán números complejos. Se puede suponer que estos valores se señalan en otro ejemplar de plano complejo, del mismo modo que en el caso de las funciones reales de la variable real los valores de la variable independiente se señalan en una recta numérica (eje de abscisas), y los valores de la función, en otra (eje de ordenadas).

La definición de función continua que conoce el lector por el curso de análisis matemático, se generaliza también para la función de la variable compleja, donde en el enunciado de la definición se deben sustituir los valores absolutos por los módulos.

Precisando, la función compleja $f(x)$ de la variable compleja x se llama *continua en el punto* x_0, si para cualquier número real positivo ε se puede elegir un número real positivo δ tal que se cumpla la desigualdad

$$|f(x_0+h)-f(x_0)|<\varepsilon$$

para cualquier incremento h (por lo general, complejo) cuyo módulo satisfaga a la desigualdad $|h|<\delta$. La función $f(x)$ se llama *continua*, si es continua en todos los puntos x_0 en que está definida la función.

Un polinomio $f(x)$ representa una función continua de la variable compleja x.

Se podría efectuar la demostración de este teorema del mismo modo que se hace en el curso de análisis matemático, o sea, demostrando que la suma y el producto de funciones continuas también son continuas y observando que una función que constantemente es igual a un mismo número complejo, es continua. Sin embargo, aquí procederemos de otro modo.

Demostraremos primero un caso particular del teorema: la continuidad de $f(x)$ en el punto $x_0 = 0$, suponiendo que el término independiente del polinomio $f(x)$ es igual a cero. En otras palabras, demostraremos el siguiente lema (en lugar de h se escribirá x):

Lema 1. *Si el término independiente del polinomio $f(x)$ es igual a cero:*

$$f(x) = a_0x^n a_1x^{n-1} + \dots + a_{n-1}x,$$

o sea, $f(0) = 0$, entonces para cualquier $\varepsilon > 0$ existe un número $\delta > 0$ tal, que $|f(x)| < \varepsilon$ para todos los x que satisfacen a la condición $|x| < \delta$.

En efecto, sea

$$A = \text{máx}(|a_0|, |a_1|, \dots, |a_{n-1}|).$$

Sea dado el número ε. Demostremos que el número

$$\delta = \frac{\varepsilon}{A+\varepsilon}, \tag{1}$$

satisface a las condiciones que se piden.

En efecto,

$$|f(x)| \leqslant |a_0|\,|x|^n + |a_1|\,|x|^{n-1} + \dots + |a_{n-1}|\,|x| \leqslant$$
$$\leqslant A(|x|^n + |x|^{n-1} + \dots + |x|),$$

o sea,

$$|f(x)| \leqslant A\frac{|x| - |x|^{n+1}}{1-|x|}.$$

Como $|x| < \delta$ y como, en virtud de (1), $\delta < 1$, se tiene:

$$\frac{|x| - |x|^{n+1}}{1-|x|} < \frac{|x|}{1-|x|},$$

por lo cual,

$$|f(x)| < \frac{A|x|}{1-|x|} < \frac{A\delta}{1-\delta} = \frac{A\dfrac{\varepsilon}{A+\varepsilon}}{1-\dfrac{\varepsilon}{A+\varepsilon}} = \varepsilon,$$

como se quería demostrar.

Deduzcamos ahora la fórmula que sigue. Sea dado un polinomio

$$f(x) = a_0x^n + a_1x^{n-1} + \dots + a_{n-1}x + a_n$$

con cualesquiera coeficientes complejos. Sustituyamos x por la suma $x+h$, donde h es otra indeterminada. Desarrollando en el primer miembro cada una de las potencias $(x+h)^k$, $k \leqslant n$, según la fórmula del binomio, y reuniendo todos los términos con iguales potencias de h, se obtiene la igualdad

$$f(x+h)=f(x)+hf'(x)+\frac{h^2}{2!}f''(x)+\ldots+\frac{h^n}{n!}f^{(n)}(x),$$

que el lector fácilmente puede comprobar, o sea, la *fórmula de Taylor*, que proporciona el desarrollo de $f(x+h)$ en potencias del «incremento» h.

La continuidad de un polinomio arbitrario $f(x)$ en cualquier punto x_0 se demuestra ahora del modo siguiente. En virtud de la fórmula de Taylor,

$$f(x_0+h)-f(x_0)=c_1h+c_2h^2+\ldots+c_nh^n=\varphi(h),$$

donde

$$c_1=f'(x_0),\ c_2=\frac{1}{2!}f''(x_0),\ \ldots,\ c_n=\frac{1}{n!}f^{(n)}(x_0).$$

El polinomio $\varphi(h)$ en la indeterminada h es un polinomio sin término independiente y, por esto, en virtud del lema 1, para cualquier $\varepsilon>0$ existe un $\delta>0$ tal que $|\varphi(h)|<\varepsilon$, o sea,

$$|f(x_0+h)-f(x_0)|<\varepsilon,$$

para $h<\delta$, que es lo que se quería demostrar.

De la desigualdad

$$||f(x_0+h)|-|f(x_0)||\leqslant|f(x_0+h)-f(x_0)|,$$

basada en la fórmula (13) del § 18, y de la continuidad de un polinomio, que acabamos de demostrar, se deduce la *continuidad de módulo* $|f(x)|$ del polinomio $f(x)$; es evidente que este módulo es una función real no negativa de la variable compleja x.

Ahora se demostrarán unos lemas que se empleara en la demostración del teorema fundamental.

Lema sobre el módulo del término superior. *Dado un polinomio de n-ésimo grado,* $n\geqslant 1$,

$$f(x)=a_0x^n+a_1x^{n-1}+a_2x^{n-2}+\ldots+a_n$$

de coeficientes complejos arbitrarios, si k *es un número real positivo cualquiera, para los valores de la indeterminada* x*, cuyos módulos son suficientemente grandes, se verifica la desigualdad*

$$|a_0x^n|>k|a_1x^{n-1}+a_2x^{n-2}+\ldots+a_n|, \tag{2}$$

es decir, el módulo del término superior es mayor que el módulo de la suma de todos los demás términos, y además, una cantidad arbitraria de veces.

En efecto, sea A el máximo de los módulos de los coeficientes $a_1, a_2, \ldots, a_n$:

$$A = \text{máx}(|a_1|, |a_2|, \ldots, |a_n|).$$

Entonces

$$|a_1x^{n-1} + a_2x^{n-2} + \ldots + a_n| \leqslant |a_1| |x|^{n-1} + |a_2| |x|^{n-2} + \ldots$$
$$\ldots + |a_n| \leqslant A(|x|^{n-1} + |x|^{n-2} + \ldots + 1) = A\frac{|x|^n - 1}{|x| - 1}.$$

(véase en el § 18 las propiedades de los módulos de la suma y el producto de números complejos).

Suponiendo $|x| > 1$, se obtiene:

$$\frac{|x|^n - 1}{|x| - 1} < \frac{|x|^n}{|x| - 1},$$

de donde

$$|a_1x^{n-1} + a_2x^{n-2} + \ldots + a_n| < A\frac{|x|^n}{|x| - 1}.$$

Por lo tanto, se cumplirá la desigualdad (2) si x, además de satisfacer a la condición $|x| > 1$, satisface también a la desigualdad

$$kA\frac{|x|^n}{|x| - 1} \leqslant |a_0x^n| = |a_0| |x|^n,$$

o sea, si

$$|x| \geqslant \frac{kA}{|a_0|} + 1. \tag{3}$$

Como el segundo miembro de la desigualdad (3) es mayor que 1, se puede afirmar que para los valores de x que satisfagan a esta desigualdad, se cumple la desigualdad (2), la cual demuestra el lema.

Lema sobre el crecimiento del módulo de un polinomio. *Para todo polinomio $f(x)$ de coeficientes complejos, cuyo grado no sea menor que la unidad, y para cualquier número real positivo M arbitrariamente grande, se puede elegir un número real positivo N tal, que $|f(x)| > M$ para $|x| > N$.*

Sea

$$f(x) = a_0x^n + a_1x^{n-1} + \ldots + a_n.$$

En virtud de la fórmula (11) del § 18,

$$|f(x)| = |a_0x^n + (a_1x^{n-1} + \ldots + a_n)| \geqslant$$
$$\geqslant |a_0x^n| - |a_1x^{n-1} + \ldots + a_n|. \tag{4}$$

Apliquemos el lema sobre el módulo del término superior. Suponiendo $k = 2$, existe un número N_1 tal que

$$|a_0x^n| > 2|a_1x^{n-1} + \ldots + a_n|,$$

para $|x| > N_1$.

De aquí que

$$|a_1x^{n-1}+\ldots+a_n|<\frac{1}{2}|a_0x^n|,$$

o sea, en virtud de (4),

$$|f(x)|>|a_0x^n|-\frac{1}{2}|a_0x^n|=\frac{1}{2}|a_0x^n|.$$

El segundo miembro de esta desigualdad será mayor que M para

$$|x|>N_2=\sqrt[n]{\frac{2M}{|a_0|}}.$$

Por lo tanto, $|f(x)|>M$ para $|x|>N=\text{máx}\,(N_1,\ N_2)$.

Se puede aclarar el significado de este lema mediante la siguiente ilustración geométrica, que se empleará a menudo en este párrafo. Supongamos que por cada punto x_0 del plano complejo se traza una perpendicular a este plano de longitud (referida a la unidad de medida elegida) igual al módulo del valor del polinomio $f(x)$ en este punto, o sea, igual a $|f(x_0)|$. En virtud del teorema de la continuidad del módulo de un polinomio, demostrado anteriormente, los extremos de estas perpendiculares formarán una superficie alabeada continua, situada sobre el plano complejo. El lema sobre el crecimiento del módulo de un polinomio muestra que al aumentar $|x_0|$, esta superficie se aleja más y más del plano complejo, aunque, naturalmente, este alejamiento no es monótono. La fig. 8 representa esquemáticamente la línea de intersección de esta superficie con un plano perpendicular al plano complejo y pasa por el punto O.

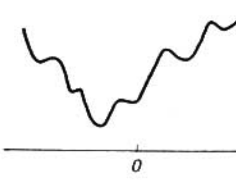


Fig. 8.

En la demostración, el papel fundamental lo desempeña el siguiente lema:

Lema de D'Alembert. *Si para $x=x_0$ el polinomio $f(x)$ de grado n, $n\geqslant 1$,* no se anula, $f(x_0)\neq 0$ y, por lo tanto, $|f(x_0)|>0$, *se puede hallar un incremento h, generalmente complejo*, tal que

$$|f(x_0+h|<|f(x_0)|.$$

Según la fórmula de Taylor, siendo por ahora arbitrario el incremento h, se tiene:

$$f(x_0+h)=f(x_0)+hf'(x_0)+\frac{h^2}{2!}f''(x_0)+\ldots+\frac{h^n}{n!}f^{(n)}(x_0).$$

Por la condición, x_0 no es raíz de $f(x)$. Sin embargo, este número puede ser eventualmente, raíz de $f'(x)$, y también puede ocurrir que

sea raíz de ciertas derivadas posteriores. Supongamos que la k-ésima derivada ($k \geqslant 1$) es la primera que no tiene a x_0 por raíz, o sea, que

$$f'(x_0) = f''(x_0) = \ldots = f^{(k-1)}(x_0) = 0, \quad f^{(k)}(x_0) \neq 0.$$

Tal k existe, puesto que si a_0 es el coeficiente superior del polinomio $f(x)$, entonces

$$f^{(n)}(x_0) = n!\, a_0 \neq 0.$$

Por lo tanto

$$f(x_0 + h) = f(x_0) + \frac{h^k}{k!} f^{(k)}(x_0) + \\ + \frac{h^{k+1}}{(k+1)!} f^{(k+1)}(x_0) + \ldots + \frac{h^n}{n!} f^{(n)}(x_0).$$

Algunos de los números $f^{(k+1)}(x_0), \ldots, f^{(n-1)}(x_0)$ también pueden ser iguales a cero. Pero esto no importa.

Dividiendo ambos miembros de esta igualdad por $f(x_0)$, que por la condición es diferente de cero, e introduciendo la notación

$$c_j = \frac{f^{(j)}(x_0)}{j!\, f(x_0)}, \quad j = k, \ k+1, \ \ldots, \ n,$$

se obtiene:

$$\frac{f(x_0 + h)}{f(x_0)} = 1 + c_k h^k + c_{k+1} h^{k+1} + \ldots + c_n h^n$$

o, puesto que $c_k \neq 0$,

$$\frac{f(x_0 + h)}{f(x_0)} = (1 + c_k h^k) + c_k h^k \left(\frac{c_{k+1}}{c_k} h + \ldots + \frac{c_n}{c_k} h^{n-k} \right).$$

Pasando a los módulos se obtiene:

$$\left| \frac{f(x_0 + h)}{f(x_0)} \right| \leqslant |1 + c_k h^k| + |c_k h^k| \left| \frac{c_{k+1}}{c_k} h + \ldots + \frac{c_n}{c_k} h^{n-k} \right|. \quad (5)$$

Hasta ahora no hemos hecho ninguna suposición sobre el incremento h. Ahora vamos a **elegir** h; además, elegiremos su módulo y su argumento por separado. El módulo de h se elegirá del modo siguiente. Como

$$\frac{c_{k+1}}{c_k} h + \ldots + \frac{c_n}{c_k} h^{n-k}$$

es un polinomio en h sin término independiente, en virtud del lema 1 $\left(\text{suponiendo } \varepsilon = \frac{1}{2}\right)$, se puede hallar un δ_1 tal que

$$\left| \frac{c_{k+1}}{c_k} h + \ldots + \frac{c_n}{c_k} h^{n-k} \right| < \frac{1}{2}, \quad (6)$$

para $|h| < \delta_1$.

Por otra parte,

$$|c_k h^k| < 1, \quad (7)$$

para

$$|h| < \delta_2 = \sqrt[k]{|c_k|^{-1}}.$$

Supongamos que el módulo de h se ha elegido de acuerdo a la desigualdad

$$|h| < \text{mín}(\delta_1, \ \delta_2). \tag{8}$$

Entonces, en virtud de (6), la desigualdad (5) se convierte en la desigualdad estricta

$$\left|\frac{f(x_0+h)}{f(x_0)}\right| < |1+c_kh^k| + \frac{1}{2}|c_kh^k|; \tag{9}$$

un poco más adelante utilizaremos la condición (7).

Para la elección del argumento de h, exigiremos que el número c_kh_k sea real y negativo. En otras palabras,

$$\arg(c_kh^k) = \arg c_k + k \arg h = \pi,$$

de donde

$$\arg h = \frac{\pi - \arg c_k}{k}. \tag{10}$$

Con esta elección de h, el número c_kh^k se diferenciará de su valor absoluto en el signo,

$$c_kh^k = -|c_kh^k|;$$

por consiguiente, aplicando la desigualdad (7),

$$|1+c_kh^k| = |1-|c_kh^k|| = 1-|c_kh^k|.$$

En consecuencia, eligiendo h de acuerdo a las condiciones (8) y (10), la desigualdad (9) toma la forma

$$\left|\frac{f(x_0+h)}{f(x_0)}\right| < 1-|c_kh^k| + \frac{1}{2}|c_kh^k| = 1-\frac{1}{2}|c_kh^k|,$$

o sea,

$$\left|\frac{f(x_0+h)}{f(x_0)}\right| = \frac{|f(x_0+h)|}{|f(x_0)|} < 1,$$

de donde resulta

$$|f(x_0+h)| < |f(x_0)|,$$

lo que demuestra el lema de D'Alembert.

Mediante la ilustración geométrica que se dio anteriormente, se puede aclarar el lema de D'Alembert del modo siguiente. Supongamos que $|f(x_0)| > 0$. Esto significa que la longitud de la perpendicular al plano complejo, trazada por el punto x_0, es diferente de cero. Entonces, según el lema de D'Alembert, se puede hallar un punto $x_1 = x_0 + h$ tal que $|f(x_1)| < |f(x_0)|$, o sea, la perpendicular en el punto x_1 será más corta que en el punto x_0, y, por consiguiente, la superficie formada por los extremos de las perpen-

diculares estará en este punto nuevo un poco más cerca del plano complejo. Como se ve por la demostración del lema, el módulo de h se puede suponer lo más pequeño que se desee, o sea, el punto x_1 se puede elegir cuanto más cerca de x_0 se quiera; sin embargo, no aplicaremos a continuación esta observación.

Es evidente que son raíces del polinomio $f(x)$ los números complejos (o sea, los puntos del plano complejo) en los que la superficie formada por los extremos de las perpendiculares está en contacto con este plano. Basándose solamente en el lema de D'Alembert no se puede demostrar la existencia de tales puntos. En efecto, aplicando este lema se puede hallar una sucesión indefinida de puntos $x_0, x_1, x_2, \ldots$, tal que

$$|f(x_0)| > |f(x_1)| > |f(x_2)| > \ldots \qquad (11)$$

Sin embargo, de aquí no se deduce la existencia de un punto $\bar{x}$ tal que $f(\bar{x}) = 0$; pues una sucesión decreciente de números reales positivos (11) no tiende necesariamente a cero.

El examen ulterior se basará en un teorema de la teoría de las funciones de la variable compleja que generaliza el teorema de Weierstrass, conocido por el lector en el curso de análisis matemático. Este se refiere a las funciones reales de la variable compleja, es decir, a las funciones de la variable compleja que solamente toman valores reales; un ejemplo de tales funciones es el módulo de un polinomio. En el enunciado de este teorema, para simplificar, se hablará del *círculo cerrado E*, entendiendo por esto un círculo del plano complejo al que se le han añadido todos los puntos de su contorno.

Si una función real $g(x)$ de la variable compleja x es continua en todos los puntos de un círculo cerrado E, en éste existe un punto x_0 tal, que para todos los puntos x de E se verifica la desigualdad $g(x) \geqslant g(x_0)$. Por consiguiente, el valor mínimo de $g(x)$ en el círculo E se alcanza en el punto x_0.

La demostración de este teorema se puede hallar en todos los cursos de teoría de las funciones de la variable compleja, por lo que aquí no se expondrá.

Aclararemos geométricamente este teorema mediante la ilustración empleada anteriormente, limitándonos al caso en que la función $g(x)$ sea no negativa en todos los puntos del círculo E (solamente este caso presenta interés). Tracemos por cada punto x_0 del círculo E una perpendicular de longitud $g(x_0)$. Los extremos de estas perpendiculares formarán un trozo de una superficie alabeada continua, y como el círculo E es cerrado, geométricamente está suficientemente claro que para esta superficie existe un mínimo. Naturalmente, esta ilustración no sustituye a la demostración del teorema.

Ahora podemos pasar a la **demostración directa del teorema fundamental.** Sea dado un polinomio $f(x)$ de grado n, $n \geqslant 1$. Resulta evidente que si su término independiente es a_n, se tiene: $f(0) = a_n$. Apliquemos el lema sobre el crecimiento del módulo de un polinomio, suponiendo $M = |f(0)| = |a_n|$. Por consiguiente, existe un N tal que $|f(x)| > |f(0)|$ para $|x| > N$. Se comprende que la generalización del teorema de Weierstrass indicada anteriormente es aplicable a la función $|f(x)|$ para cualquier círculo cerrado E elegido. Tomemos por E el círculo cerrado, limitado por la circunferencia de radio N con centro en el punto 0. Supongamos que en el círculo E, la función $|f(x)|$ alcanza el mínimo en el punto x_0; entonces, en particular, se tiene: $|f(x_0)| \leqslant |f(0)|$.

Fácilmente se observa que *en todo el plano complejo la función* $|f(x)|$ *alcanza el mínimo en el punto* x_0: si el punto x' está situado fuera de E, se tiene $|x'| > N$, por lo cual,

$$|f(x')| > |f(0)| \geqslant |f(x_0)|.$$

Finalmente, de aquí se deduce que $f(x_0) = 0$, o sea, que x_0 *es raíz de* $f(x)$; si fuese $f(x_0) \neq 0$, entonces, por el lema de D'Alembert, existiría un punto x_1 tal que $|f(x_1)| < |f(x_0)|$; sin embargo, esto contradice a la propiedad del punto x_0 que acabamos de establecer.

En el § 55 se dará otra demostración del teorema fundamental.

§ 24. Consecuencias del teorema fundamental

Sea dado un polinomio de n-ésimo grado, $n \geqslant 1$,

$$f(x) = a_0x^n + a_1x^{n-1} + \ldots + a_{n-1}x + a_n \tag{1}$$

con cualesquiera coeficientes complejos. De nuevo lo consideramos como una expresión algebraica formal, determinada completamente por el conjunto de sus coeficientes. El teorema fundamental de existencia de la raíz, demostrado en el párrafo anterior, permite afirmar la existencia de una raíz α_1 de $f(x)$, que puede ser real o compleja. Por lo tanto, el polinomio $f(x)$ se puede descomponer en la forma

$$f(x) = (x - \alpha_1)\,\varphi(x).$$

Los coeficientes del polinomio $\varphi(x)$ son de nuevo números reales o complejos y, por consiguiente, $\varphi(x)$ tiene una raíz α_2, de donde,

$$f(x) = (x - \alpha_1)(x - \alpha_2)\,\varphi(x).$$

Continuando de este modo, después de un número finito de operaciones, obtendremos la **descomposición del polinomio** $f(x)$ **de n-ésimo grado en un producto de n factores lineales,**

$$f(x) = a_0(x - \alpha_1)(x - \alpha_2)\ldots(x - \alpha_n). \tag{2}$$

La causa de la aparición del coeficiente a_0 es la siguiente: si en el segundo miembro de la expresión (2) figurase cierto coeficiente b, después de abrir paréntesis el término superior del polinomio $f(x)$ tendría la forma bx^n, mientras que éste es igual a a_0x^n, en virtud de (1). Por esto, $b = a_0$.

La descomposición (2) del polinomio $f(x)$ es la única descomposición de este tipo, salvo el orden de los factores.

En efecto, supongamos que haya otra descomposición

$$f(x) = a_0(x-\beta_1)(x-\beta_2)\ldots(x-\beta_n). \tag{3}$$

De (2) y (3) se deduce la igualdad

$$(x-\alpha_1)(x-\alpha_2)\ldots(x-\alpha_n) = (x-\beta_1)(x-\beta_2)\ldots(x-\beta_n). \tag{4}$$

Si la raíz α_i fuese distinta de todas las β_j, $j = 1, 2, \ldots, n$, sustituyendo en (4) α_i en lugar de la indeterminada, obtendríamos cero en el primer miembro, mientras que en el segundo miembro, un número diferente de cero. Por lo tanto, *toda raiz α_i es igual a cierta raíz β_j, y viceversa.*

De aquí todavía no se deduce la identidad de las descomposiciones (2) y (3). En efecto, entre las raíces α_i, $i = 1, 2, \ldots, n$, puede haber iguales entre sí. Supongamos, por ejemplo, que s de estas raíces son iguales a α_1 y que, por otra parte, entre las raíces β_j, $j = 1, 2, \ldots, n$, hay exactamente t iguales a la raíz α_1. Se necesita demostrar que $s = t$.

Como el grado de un producto de polinomios es igual a la suma de los grados de los factores, el producto de dos polinomios diferentes de cero, no puede ser igual a cero. De aquí se deduce que si dos productos de polinomios son iguales entre sí, *ambos miembros de la igualdad se pueden simplificar por el factor común*: si

$$f(x)\varphi(x) = g(x)\varphi(x)$$

y $\varphi(x) \neq 0$, de la igualdad

$$[f(x) - g(x)]\varphi(x) = 0$$

se deduce que

$$f(x) - g(x) = 0,$$

o sea,

$$f(x) = g(x).$$

Apliquemos esto a la igualdad (4). Si, por ejemplo, fuese $s > t$, simplificando ambos miembros de la igualdad (4) por el factor $(x - \alpha_1)^t$, llegaríamos a una igualdad cuyo primer miembro contendría el factor $x - \alpha_1$, mientras que el segundo miembro, no lo contendría. Sin embargo, antes se demostró que esto conduce a una contradicción. Por lo tanto, la unicidad de la descomposición (2) del polinomio $f(x)$ queda demostrada.

Reuniendo todos los factores equivalentes, se puede escribir la descomposición (2) en la forma*

$$f(x) = a_0 (x - \alpha_1)^{k_1} (x - \alpha_2)^{k_2} \dots (x - \alpha_l)^{k_l}, \tag{5}$$

donde

$$k_1 + k_2 + \dots + k_l = n.$$

Aquí se supone que entre las raíces $\alpha_1, \alpha_2, \dots, \alpha_l$ *ya no hay iguales.*

Demostremos que en (5), el *número* k_i, $i = 1, 2, \dots, l$, *es el orden de multiplicidad de la raíz* α_i *del polinomio* $f(x)$. En efecto, si este orden es igual a s_i, entonces, $k_i \leqslant s_i$. Sin embargo, supongamos que $k_i < s_i$. En virtud de la definición del orden de multiplicidad de la raíz, para $f(x)$ subsiste la descomposición

$$f(x) = (x - \alpha_i)^{s_i} \varphi(x).$$

Sustituyendo en esta descomposición el factor $\varphi(x)$ por su descomposición en factores lineales, obtendríamos una descomposición de $f(x)$ en factores lineales, diversa de la descomposición (2), o sea, llegaríamos a una contradicción con la unicidad de esta descomposición, demostrada anteriormente.

Por lo tanto, hemos demostrado el siguiente resultado importante:

Todo polinomio $f(x)$ *de grado* n, $n \geqslant 1$, *de cualesquiera coeficientes numéricos, tiene* n *raíces, contando cada una de las raíces tantas veces como sea su orden de multiplicidad.*

Obsérvese que nuestro teorema subsiste también para $n = 0$, puesto que un polinomio de grado cero, no tiene raíces. Este teorema no se cumple **solamente** para el polinomio 0, el cual no tiene grado alguno y es igual a cero para cualquier valor de x. Esta última observación se utilizará para la demostración del siguiente **teorema:**

Si los polinomios $f(x)$ y $g(x)$ *de grado no superior a* n, *toman valores iguales para más de* n *valores de la indeterminada, entonces* $f(x) = g(x)$.

En efecto, en nuestras condiciones, el polinomio $f(x) - g(x)$ tiene más de n raíces, y como es de grado no superior a n, se cumple la igualdad $f(x) - g(x) = 0$.

Por lo tanto, teniendo en cuenta que hay una infinidad de diversos números, se puede afirmar que *para dos polinomios* $f(x)$ y $g(x)$ *cualesquiera existen tales valores* c *de la indeterminada* x *que* $f(c) \neq g(c)$. Tales c no sólo se pueden hallar entre los números complejos, sino también entre los números reales, entre los racionales e incluso entre los números enteros.

Por consiguiente, **dos polinomios de coeficientes numéricos que tienen diferentes coeficientes, aunque sólo sea en una potencia de la indeterminada** x, *son diversas funciones complejas de la variable compleja* x. Con esto, queda por fin demostrada *para los polinomios*

---

* Se llama descomposición factorial del polinomio $f(x)$. (*Nota del T.*)

de coeficientes numéricos la equivalencia de las dos definiciones de igualdad de los polinomios (la algebraica y la teórico-funcional), indicadas en el § 20.

El teorema demostrado anteriormente permite afirmar que un *polinomio de grado no mayor que n se determina completamente por sus valores para cualesquiera valores distintos de la indeterminada, tomados en cantidad mayor que n.* ¿Pueden ser arbitrarios estos valores del polinomio? Si se supone que se dan los valores del polinomio para $n+1$ valores diversos de la indeterminada, la respuesta es positiva: *siempre existe un polinomio de grado no mayor que n, que tome unos valores prefijados para $n+1$ diversos valores dados de la indeterminada.*

En efecto, supongamos que se necesita hallar un polinomio de grado no mayor que n tal, que para los diferentes valores de la indeterminada $a_1, a_2, \ldots, a_{n+1}$, tome respectivamente los valores $c_1, c_2, \ldots, c_{n+1}$. Este polinomio es:

$$f(x)=\sum_{i=1}^{n+1}\frac{c_i(x-a_1)\ldots(x-a_{i-1})(x-a_{i+1})\ldots(x-a_{n+1})}{(a_i-a_1)\ldots(a_i-a_{i-1})(a_i-a_{i+1})\ldots(a_i-a_{n+1})}. \qquad (6)$$

En efecto, su grado no es mayor que n, y el valor $f(a_i)$ es igual a c_i.

La fórmula (6) se denomina *fórmula de interpolación de Lagrange.* La denominación «de interpolación» se debe a que, conociendo los valores del polinomio en $n+1$ puntos, se pueden hallar por esta fórmula sus valores en cualesquiera otros puntos.

Fórmulas de Vieta. Sea dado un polinomio $f(x)$ de grado n cuyo coeficiente superior es igual a 1:

$$f(x)=x^n+a_1x^{n-1}+a_2x^{n-2}+\ldots+a_{n-1}x+a_n, \qquad (7)$$

y sean $\alpha_1, \alpha_2, \ldots, \alpha_n$ sus raíces*. Entonces la descomposición factorial es

$$f(x)=(x-\alpha_1)(x-\alpha_2)\ldots(x-\alpha_n).$$

Multiplicando los paréntesis que figuran en el segundo miembro, reduciendo luego los términos semejantes y comparando los coeficientes obtenidos con los coeficientes de (7), se obtienen las siguientes igualdades, denominadas *fórmulas de Vieta*, que expresan los coeficientes del polinomio mediante sus raíces:

$$\begin{aligned}
a_1&=-(\alpha_1+\alpha_2+\ldots+\alpha_n),\\
a_2&=\alpha_1\alpha_2+\alpha_1\alpha_3+\ldots+\alpha_1\alpha_n+\alpha_2\alpha_3+\ldots+\alpha_{n-1}\alpha_n,\\
a_3&=-(\alpha_1\alpha_2\alpha_3+\alpha_1\alpha_2\alpha_4+\ldots+\alpha_{n-2}\alpha_{n-1}\alpha_n),\\
&\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\\
a_{n-1}&=(-1)^{n-1}(\alpha_1\alpha_2\ldots\alpha_{n-1}+\alpha_1\alpha_2\ldots\alpha_{n-2}\alpha_n+\ldots+\alpha_2\alpha_3\ldots\alpha_n),\\
a_n&=(-1)^n\alpha_1\alpha_2\ldots\alpha_n.
\end{aligned}$$

* Aquí se toma cada raíz múltiple el número respectivo de veces.

Por lo tanto, en el segundo miembro de la k-ésima igualdad, $k = 1, 2, \ldots, n$, figura una suma de todos los productos posibles de k raíces, tomadas con el signo más o menos, según que k sea par o impar.

Para $n = 2$, estas fórmulas se convierten en las relaciones entre las raíces y los coeficientes de un polinomio cuadrático, conocidas por el álgebra elemental. Para $n = 3$, o sea, para un polinomio cúbico, estas fórmulas toman la forma:

$$a_1 = -(\alpha_1 + \alpha_2 + \alpha_3), \quad a_2 = \alpha_1\alpha_2 + \alpha_1\alpha_3 + \alpha_2\alpha_3, \quad a_3 = -\alpha_1\alpha_2\alpha_3.$$

Las fórmulas de Vieta facilitan la escritura del polinomio, conocidas sus raíces. Así, pues, hallemos el polinomio $f(x)$ de cuarto grado, de modo que los números 5 y -2 sean raíces simples y el número 3, raíz múltiple de orden dos. Obtenemos:

$$a_1 = -(5-2+3+3) = -9,$$
$$a_2 = 5\cdot(-2)+5\cdot 3+5\cdot 3+(-2)\cdot 3+(-2)\cdot 3+3\cdot 3 = 17,$$
$$a_3 = -[5\cdot(-2)\cdot 3+5\cdot(-2)\cdot 3+5\cdot 3\cdot 3+(-2)\cdot 3\cdot 3] = 33,$$
$$a_4 = 5\cdot(-2)\cdot 3\cdot 3 = -90,$$

por lo cual

$$f(x) = x^4 - 9x^3 + 17x^2 + 33x - 90.$$

Si el coeficiente superior a_0 del polinomio $f(x)$ es diferente de 1, para la aplicación de las fórmulas de Vieta es necesario dividir primero todos los coeficientes por a_0, pues esto no influye en las raíces del polinomio. En este caso, las fórmulas de Vieta dan las expresiones para las razones de todos los coeficientes al coeficiente superior.

Polinomios de coeficientes reales. Ahora se deducirán algunas consecuencias del teorema fundamental del álgebra de los números complejos, referentes a los polinomios de coeficientes reales. Precisamente en estas consecuencias está basada la importancia exclusiva del teorema fundamental antes mencionado.

Supongamos que el polinomio de coeficientes reales

$$f(x) = a_0x^n + a_1x^{n-1} + \ldots + a_{n-1} + a_n$$

tiene la raíz imaginaria α, o sea, que

$$a_0\alpha^n + a_1\alpha^{n-1} + \ldots + a_{n-1}\alpha + a_n = 0.$$

Ya sabemos que no se infringe la última igualdad al sustituir todos los números por los conjugados. Sin embargo, siendo reales todos los coeficientes $a_0, a_1, \ldots, a_{n-1}, a_n$, inclusive el número 0 que figura en el segundo miembro, éstos no se alteran en esta sustitución, obteniéndose la igualdad

$$a_0\overline{\alpha}^n + a_1\overline{\alpha}^{n-1} + \ldots + a_{n-1}\overline{\alpha} + a_n = 0,$$

o sea,

$$f(\bar{\alpha}) = 0.$$

Por lo tanto, *si un número imaginario* α *es una raíz de un polinomio* $f(x)$ *de coeficientes reales, el número conjugado* $\bar{\alpha}$ *también es una raíz de* $f(x)$.

Por consiguiente, el polinomio $f(x)$ es divisible por el trinomio cuadrático

$$\varphi(x) = (x - \alpha)(x - \bar{\alpha}) = x^2 - (\alpha + \bar{\alpha})x + \alpha\bar{\alpha}, \tag{8}$$

cuyos coeficientes, como ya sabemos por el § 18, son reales. Aplicando esto, demostremos que las *raíces* α *y* $\bar{\alpha}$ *del polinomio* $f(x)$ *son de un mismo orden de multiplicidad.*

En efecto, supongamos que los órdenes de multiplicidad de estas raíces son k y l, y que, por ejemplo, $k > l$. Entonces $f(x)$ es divisible por la l-ésima potencia del polinomio $\varphi(x)$.

$$f(x) = \varphi^l(x)\, q(x).$$

El polinomio $q(x)$, como cociente de dos polinomios de coeficientes reales, también tiene coeficientes reales, pero, en contra de lo demostrado anteriormente, el número α es raíz de éste de orden $(k - l)$, mientras que el número $\bar{\alpha}$ no es raíz. De aquí se deduce que $k = l$.

Por lo tanto, ahora se puede decir que las *raíces imaginarias de todo polinomio de coeficientes reales son conjugadas a pares.* De aquí y de la unicidad de las descomposiciones de la forma (2), demostrada anteriormente, se deduce el siguiente resultado final:

Todo polinomio $f(x)$ *de coeficientes reales se descompone de modo único (salvo el orden de los factores) en forma de un producto de su coeficiente superior* a_0 *y de unos cuantos polinomios de coeficientes reales, unos de los cuales son lineales de la forma* $x - \alpha$, *correspondientes a sus raíces reales, y otros, son cuadrados de la forma* (8), *correspondientes a los pares, de sus raíces imaginarias conjugadas.*

Para lo que sigue, es conveniente subrayar que entre los polinomios de coeficientes reales con el coeficiente superior 1, solamente los polinomios lineales de la forma $x - \alpha$ y los cuadrados de la forma (8), no se descomponen en factores de menor grado o, como diremos, son *irreducibles.*

§ 25. Fracciones racionales

En el curso de análisis matemático, además de las funciones racionales enteras, llamadas polinomios, se estudian también las *funciones racionales fraccionarias*; éstas son los cocientes $\frac{f(x)}{g(x)}$ de dos funciones racionales enteras, donde $g(x) \neq 0$. Con estas funciones se efectúan operacio-

nes algebraicas según las mismas leyes con que se opera con los números racionales, o sea, como con quebrados de numeradores y denominadores enteros. La igualdad de dos funciones racionales fraccionarias o, como en adelante se dirá, de dos *fracciones racionales*, se entenderá también en el mismo sentido que la igualdad de quebrados en la aritmética elemental. Para precisar, consideraremos las fracciones racionales con *coeficientes* reales; el lector observará sin dificultad que todo el contenido del presente párrafo se puede trasladar casi palabra por palabra, al caso de fracciones racionales con coeficientes complejos.

Una fracción racional se llama *irreducible*, si su numerador es primo con su denominador.

Toda fracción racional es igual a una fracción irreducible, determinada unívocamente salvo un factor de grado cero, que es común para el numerador y denominador.

En efecto, cualquier fracción racional se puede simplificar por el máximo común divisor de su numerador y denominador, después de lo cual resulta una fracción irreducible igual a la dada. Después, si las fracciones irreducibles $\frac{f(x)}{g(x)}$ y $\frac{\varphi(x)}{\psi(x)}$ son iguales entre sí, o sea, si

$$f(x)\,\psi(x) = g(x)\,\varphi(x), \tag{1}$$

como $f(x)$ y $g(x)$ son primos entre sí, por la propiedad b) del § 21 se deduce que $\varphi(x)$ es divisible por $f(x)$; y como $\varphi(x)$ y $\psi(x)$ son primos entre sí, resulta que $f(x)$ es divisible por $\varphi(x)$. Por lo tanto, $f(x) = c\varphi(x)$, y de (1) se deduce que $g(x) = c\,\psi(x)$.

Una fracción racional se dice que es *propia*, si el grado del numerador es menor que el grado del denominador. Si convenimos en considerar al polinomio 0 como una fracción propia, subsiste el siguiente **teorema:**

Toda fracción racional se representa de un modo único en forma de una suma de un polinomio y una fracción propia.

En efecto, si se da una fracción racional $\frac{f(x)}{g(x)}$ y si, dividiendo el numerador por el denominador, se obtiene la igualdad

$$f(x) = g(x)\,q(x) + r(x),$$

donde el grado de $r(x)$ es menor que el grado de $g(x)$, entonces, como fácilmente se comprueba,

$$\frac{f(x)}{g(x)} = q(x) + \frac{r(x)}{g(x)}\,.$$

Si también se cumple la igualdad

$$\frac{f(x)}{g(x)} = \bar{q}(x) + \frac{\varphi(x)}{\psi(x)}\,,$$

donde el grado de $\varphi(x)$ es menor que el grado de $\psi(x)$, entonces resulta la igualdad

$$q(x)-\bar{q}(x)=\frac{\varphi(x)}{\psi(x)}-\frac{r(x)}{g(x)}=\frac{\varphi(x)g(x)-\psi(x)r(x)}{\psi(x)g(x)}.$$

Como en el primer miembro figura un polinomio, mientras que en el segundo, una fracción propia, resulta: $q(x)-\bar{q}(x)=0$ y

$$\frac{\varphi(x)}{\psi(x)}-\frac{r(x)}{g(x)}=0.$$

Las fracciones racionales propias pueden ser sometidas a un examen ulterior. Recordemos para esto que, como se ha señalado al final del párrafo anterior, son polinomios reales irreducibles los de la forma $x-\alpha$, donde α es real, y los de la forma $x^2-(\beta+\bar{\beta})x+$ $+\beta\bar{\beta}$, donde β y $\bar{\beta}$ es un par de números imaginarios conjugados. Como fácilmente se comprueba, en el caso complejo desempeñan un papel análogo los polinomios de la forma $x-\alpha$, donde α es un número complejo cualquiera.

La fracción racional propia $\frac{f(x)}{g(x)}$ se llama *simple*, si su denominador $g(x)$ es una potencia de un polinomio irreducible $p(x)$,

$$g(x)=p^k(x), \qquad k\geqslant 1,$$

y el grado del numerador $f(x)$ es menor que el grado de $p(x)$.

Subsiste también el siguiente **teorema fundamental:**

Toda fracción racional propia se descompone en una suma de fracciones simples.

Demostración. Consideremos primero la fracción racional propia $\frac{f(x)}{g(x)h(x)}$, donde los polinomios $g(x)$ y $h(x)$ son primos entre sí:

$$(g(x),\ h(x))=1.$$

Por consiguiente, en virtud del § 21, existen unos polinomios $\bar{u}(x)$ y $\bar{v}(x)$ tales que

$$g(x)\bar{u}(x)+h(x)\bar{v}(x)=1.$$

De aquí,

$$g(x)[\bar{u}(x)f(x)]+h(x)[\bar{v}(x)f(x)]=f(x). \tag{2}$$

Supongamos que dividiendo el producto $u(x)f(x)$ por $h(x)$, se obtiene un resto $u(x)$, cuyo grado es menor que el grado de $h(x)$. En este caso, la igualdad (2) se puede escribir del modo siguiente:

$$g(x)u(x)+h(x)v(x)=f(x), \tag{3}$$

donde $v(x)$ es un polinomio cuya expresión se podría haber escrito sin dificultad. Como el grado del producto $g(x)u(x)$ es menor que

el grado del producto $g(x) h(x)$ y esto mismo es cierto, según la condición, para el polinomio $f(x)$, el producto $h(x) v(x)$ será también de grado menor que $g(x) h(x)$ y, por consiguiente, el grado de $v(x)$ será menor que el grado de $g(x)$. De (3) se deduce ahora la igualdad

$$\frac{f(x)}{g(x)\,h(x)} = \frac{v(x)}{g(x)} + \frac{u(x)}{h(x)},$$

en cuyo segundo miembro figura una suma de fracciones propias.

Si al menos uno de los denominadores $g(x)$, $h(x)$ se descompone en un producto de factores primos entre sí, se puede efectuar la descomposición ulterior; continuando de este modo, se obtiene que *cualquier fracción propia se descompone en una suma de unas cuantas fracciones propias, cada una de las cuales tiene por denominador una potencia de un polinomio irreducible.* Más exactamente, dada una fracción propia $\frac{f(x)}{g(x)}$, cuyo denominador posee la descomposición en factores irreducibles

$$g(x) = p_1^{k_1}(x)\,p_2^{k_2}(x)\,\ldots\,p_l^{k_l}(x)$$

(por supuesto, siempre se puede suponer que el coeficiente superior del denominador de la fracción racional es igual a la unidad), siendo $p_i(x) \neq p_j(x)$ para $i \neq j$, se tiene

$$\frac{f(x)}{g(x)} = \frac{u_1(x)}{p_1^{k_1}(x)} + \frac{u_2(x)}{p_2^{k_2}(x)} + \ldots + \frac{u_l(x)}{p_l^{k_l}(x)};$$

todos los términos del segundo miembro de esta igualdad son fracciones propias.

No queda más que considerar una fracción propia de la forma $\frac{u(x)}{p^k(x)}$, donde $p(x)$ es un polinomio irreducible. Aplicando el algoritmo de la división con resto dividimos $u(x)$ por $p^{k-1}(x)$, luego, el resto obtenido lo dividimos por $p^{k-2}(x)$, etc.

Llegamos a las siguientes igualdades:

$$\begin{aligned} u(x) &= p^{k-1}(x)\,s_1(x) + u_1(x), \\ u_1(x) &= p^{k-2}(x)\,s_2(x) + u_2(x), \\ &\ldots\ldots\ldots\ldots\ldots\ldots \\ u_{k-2}(x) &= p(x)\,s_{k-1}(x) + u_{k-1}(x). \end{aligned}$$

Como, por la condición, el grado de $u(x)$ es menor que el grado de $p^k(x)$, y el grado de cada uno de los restos $u_i(x)$, $i = 1, 2, \ldots, k-1$, es menor que el grado del divisor correspondiente $p^{k-i}(x)$, los grados de todos los cocientes $s_1(x)$, $s_2(x)$, $\ldots$, $s_{k-1}(x)$ serán estrictamente menores que el grado del polinomio $p(x)$. El grado del último resto $u_{k-1}(x)$ será también menor que el grado de $p(x)$.

De las igualdades obtenidas, resulta:

$$u(x) = p^{k-1}(x)\,s_1(x) + p^{k-2}(x)s_2(x) + \ldots + p(x)\,s_{k-1}(x) + u_{k-1}(x).$$

De aquí, obtenemos la representación buscada de la fracción racional $\frac{u(x)}{p^k(x)}$ en forma de una suma de fracciones simples:

$$\frac{u(x)}{p^k(x)} = \frac{u_{k-1}(x)}{p^k(x)} + \frac{s_{k-1}(x)}{p^{k-1}(x)} + \ldots + \frac{s_2(x)}{p^2(x)} + \frac{s_1(x)}{p(x)}.$$

El teorema fundamental queda demostrado. Este se puede completar con el siguiente **teorema de unicidad:**

Toda fracción racional propia posee una descomposición única en suma de fracciones simples.

En efecto, supongamos que alguna fracción propia se puede expresar de dos modos en forma de una suma de fracciones simples. Restando una de estas expresiones de la otra y reduciendo los términos semejantes, se obtiene una suma de fracciones simples, idénticamente igual a cero. Supongamos que los denominadores de las fracciones simples que forman esta suma son ciertas potencias de diferentes polinomios irreducibles $p_1(x)$, $p_2(x)$, ..., $p_s(x)$ y sea $p_i^{k_i}(x)$ la potencia superior del polinomio $p_i(x)$, $i=1, 2, \ldots, s$, que figura entre los denominadores. Multiplicando ambos miembros de la igualdad considerada por el producto $p_1^{k_1-1}(x)p_2^{k_2}(x) \ldots p_s^{k_s}(x)$, todos los términos de nuestra suma, menos uno de ellos, se convierten en polinomios. En lo que se refiere al término $\frac{u(x)}{p_1^{k_1}(x)}$, éste se convierte en una fracción cuyo denominador es $p_1(x)$ y cuyo numerador es el producto $u(x)\,p_2^{k_2}(x) \ldots p_s^{k_s}(x)$. El numerador no se divide exactamente por el denominador, puesto que el polinomio $p_1(x)$ es irreducible y todos los factores del numerador son primos con él. Efectuando la división con resto se obtiene que es igual a cero la suma de un polinomio y una fracción propia diferente de cero, lo cual es imposible.

Ejemplo. Descomponer en una suma de fracciones simples la fracción propia real $\frac{f(x)}{g(x)}$, donde

$$f(x) = 2x^4 - 10x^3 + 7x^2 + 4x + 3,$$
$$g(x) = x^5 - 2x^3 + 2x^2 - 3x + 2.$$

Fácilmente se comprueba que

$$g(x) = (x+2)(x-1)^2(x^2+1),$$

donde cada uno de los polinomios $x+2$, $x-1$, x^2+1, es irreducible. De la teoría que acabamos de exponer se deduce que la descomposición buscada tiene

que tener la forma

$$\frac{f(x)}{g(x)}=\frac{A}{x+2}+\frac{B}{(x-1)^2}+\frac{C}{x-1}+\frac{Dx+E}{x^2+1}, \qquad (4)$$

donde los números A, B, C, D y E tienen que ser todavía buscados.

De (4) se deduce la igualdad

$$f(x)=A(x-1)^2(x^2+1)+B(x+2)(x^2+1)+C(x+2)(x-1)(x^2+1)+$$
$$+Dx(x+2)(x-1)^2+E(x+2)(x-1)^2. \qquad (5)$$

Identificando los coeficientes de iguales potencias de la indeterminada x en ambos miembros de la igualdad (5), obtendríamos un sistema de cinco ecuaciones lineales respecto a cinco incógnitas, A, B, C, D y E; como se deduce de lo demostrado anteriormente, este sistema tiene una solución que además, es única. Sin embargo, procederemos de otro modo.

Poniendo en la igualdad (5) $x=-2$, obtenemos la igualdad, $45A=135$, *de donde*

$$A=3. \qquad (6)$$

Poniendo luego en (5) $x=1$, obtenemos, $6B=6$, o sea

$$B=1. \qquad (7)$$

Después de esto, ponemos en la igualdad (5) $x=0$ y $x=-1$, sucesivamente. Teniendo en cuenta (6) y (7), obtenemos las ecuaciones

$$\left.\begin{aligned}-2C+2E&=1,\\ -4C-4D+4E&=-8.\end{aligned}\right\} \qquad (8)$$

De aquí,

$$D=1. \qquad (9)$$

Pongamos, finalmente, en la igualdad (5), $x=2$. Teniendo en cuenta (6), (7) y (9), llegamos a la ecuación

$$20C+4E=-52,$$

que junto con la primera de las ecuaciones (8) da

$$C=-2, \quad E=-3.$$

Por lo tanto,

$$\frac{f(x)}{g(x)}=\frac{3}{x+2}+\frac{1}{(x-1)^2}-\frac{2}{x-1}+\frac{x-3}{x^2+1}.$$

CAPITULO VI

FORMAS CUADRATICAS

§ 26. Reducción de una forma cuadrática a la forma canónica

La teoría de las formas cuadráticas tiene su origen en la geometría analítica, más precisamente, en la teoría de las curvas (y superficies) de segundo orden. Es bien sabido que la ecuación de una curva central de segundo orden en el plano, después de trasladar el origen de coordenadas rectangulares al centro de esta curva, tiene la forma

$$Ax^2 + 2Bxy + Cy^2 = D. \qquad (1)$$

Se sabe también que se puede efectuar una rotación de los ejes coordenados en un ángulo α, o sea, un cambio de las coordenadas x, y, por las coordenadas x', y':

$$\left.\begin{aligned} x &= x' \cos\alpha - y' \operatorname{sen}\alpha, \\ y &= x' \operatorname{sen}\alpha + x' \cos\alpha, \end{aligned}\right\} \qquad (2)$$

de modo que en las nuevas coordenadas la ecuación de la curva tome la forma «canónica»:

$$A'x'^2 + C'y'^2 = D; \qquad (3)$$

por consiguiente, en esta ecuación, el coeficiente del producto $x'y'$ de las indeterminadas es igual a cero. Evidentemente, la transformación de coordenadas (2) se puede interpretar como una transformación lineal de las indeterminadas (véase el § 13), la cual, además, no es degenerada, puesto que el determinante de sus coeficientes es igual a la unidad. Esta transformación se aplica al primer miembro de la ecuación (1). Por lo tanto, se puede decir que mediante la transformación lineal no degenerada (2), el primer miembro de la ecuación (1) se convierte en el primer miembro de la ecuación (3).

Numerosas aplicaciones reclamaron la elaboración de una teoría análoga para el caso en que el número de las indeterminadas, en lugar de dos, sea igual a cualquier n, y los coeficientes sean, o bien números reales, o bien números complejos cualesquiera.

Generalizando la expresión que figura en el primer miembro de la ecuación (1), llegamos al siguiente concepto.

Se llama *forma cuadrática f* en las n indeterminadas $x_1, x_2, \ldots$ $\ldots, x_n$, a una suma, en la cual cada término o es el cuadrado de una de estas indeterminadas o es el producto de dos indeterminadas diversas. Una forma cuadrática se llama *real* o *compleja* según que sus coeficientes sean números reales o cualesquiera números complejos.

Suponiendo que en la forma cuadrática f ya se ha hecho la reducción de términos semejantes, hagamos las siguientes notaciones para los coeficientes de la misma: el coeficiente de x_i^2 lo designaremos por a_{ii}, y el coeficiente del producto x_ix_j, para $i \neq j$, por $2a_{ij}$ (¡compárese con (1)!). Como $x_ix_j = x_jx_i$, el coeficiente de este producto se podría indicar también con la notación $2a_{ji}$, o sea, las notaciones introducidas suponen el cumplimiento de la igualdad:

$$a_{ji} = a_{ij}. \tag{4}$$

El término $2a_{ij}x_ix_j$ se puede escribir ahora en la forma

$$2a_{ij}x_ix_j = a_{ij}x_ix_j + a_{ji}x_jx_i,$$

y toda la forma cuadrática f, en forma de una suma de todos los términos posibles $a_{ij}x_ix_j$, donde i y j, independientemente uno de otro, toman los valores desde 1 hasta n:

$$f = \sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij}x_ix_j; \tag{5}$$

en particular, para $i = j$ resulta el término $a_{ii}x_i^2$.

Con los coeficientes a_{ij} se puede formar, evidentemente, una matriz cuadrada $A = (a_{ij})$ de orden n; ésta se llama *matriz de la forma cuadrática f*, y su rango r, *rango* de esta forma cuadrática. Si, en particular, $r = n$, o sea, si la matriz no es degenerada, la forma cuadrática f también se llama *no degenerada*. En virtud de la igualdad (4), los elementos de la matriz A, simétricos con respecto a la diagonal principal, son iguales entre sí, es decir, la matriz A es *simétrica*. Recíprocamente, para cualquier matriz simétrica A de orden n se puede indicar una forma cuadrática (5) en n indeterminadas, cuyos coeficientes son los elementos de la matriz A.

La forma cuadrática (5) puede ser escrita en otra forma, aplicando el producto de matrices rectangulares, definido en el § 14. Convengamos primero en hacer las siguientes notaciones: dada una matriz cuadrada A, o en general, una matriz rectangular, se designará por A' la matriz que se obtiene transponiendo la matriz A. Si las matrices A y B son tales que está definido su producto, entonces se cumple la igualdad:

$$(AB)' = B'A', \tag{6}$$

o sea, la *matriz transpuesta del producto es igual al producto de las matrices transpuestas de los factores, pero tomadas en orden inverso.*

En efecto, si está definido el producto AB, también estará definido el producto $B'A'$, lo que se comprueba fácilmente: el número de columnas de la matriz B' es igual al número de filas de la matriz A'. El elemento de la matriz $(AB)'$ que figura en su i-ésima fila y en su j-ésima columna, está situado en la matriz AB en la j-ésima fila e i-ésima columna. Por esto, es igual a la suma de los productos de los elementos correspondientes de la j-ésima fila de la matriz A y de la i-ésima columna de la matriz B, o sea, es igual a la suma de los productos de los elementos correspondientes de la j-ésima columna de la matriz A' y de la i-ésima fila de la matriz B'. Con esto, queda demostrada la igualdad (6).

Obsérvese que la *matriz A es simétrica cuando, y sólo cuando, ella coincide con su transpuesta*, o sea, si

$$A' = A.$$

Designemos ahora con X la columna formada por las indeterminadas:

$$X = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}.$$

X es una matriz de n filas y una columna. Transponiendo esta matriz se obtiene la matriz

$$X' = (x_1, x_2, \ldots, x_n),$$

formada por una sola fila.

La forma cuadrática (5), cuya matriz es $A = (a_{ij})$, se puede escribir ahora en forma del producto:

$$f = X'AX. \tag{7}$$

En efecto, el producto AX es una matriz formada por una columna:

$$AX = \begin{pmatrix} \sum_{j=1}^{n} a_{1j}x_j \\ \sum_{j=1}^{n} a_{2j}x_j \\ \vdots \\ \sum_{j=1}^{n} a_{nj}x_j \end{pmatrix}.$$

Multiplicando por la izquierda esta matriz por la matriz X', se obtiene una matriz», formada por una fila y una columna, que es precisamente el segundo miembro de la igualdad (5).

¿Qué ocurrirá con la forma cuadrática f si se someten las indeterminadas $x_1, x_2, \ldots, x_n$ a una trasformación lineal

$$x_i = \sum_{k=1}^{n} q_{ik} y_k, \quad i = 1, 2, \ldots, n, \tag{8}$$

de matriz $Q = (q_{ik})$? Se supone que, si la forma f es real, también tienen que ser reales los elementos de la matriz Q. Designando con Y la columna formada por las indeterminadas $y_1, y_2, \ldots, y_n$, escribamos la transformación lineal (8) en forma de una igualdad matricial:

$$X = QY. \tag{9}$$

De aquí, en virtud de (6)

$$X' = Y'Q'. \tag{10}$$

Sustituyendo (9) y (10) en la expresión (7) de la forma f, resulta:

$$f = Y' (Q'AQ) Y,$$

o

$$f = Y'BY,$$

donde

$$B = Q'AQ.$$

La matriz B es simétrica, puesto que, en virtud de la igualdad (6), que se cumple evidentemente para cualquier número de factores, y de la igualdad $A' = A$, que significa que la matriz A es simétrica, se tiene:

$$B' = Q'A'Q = Q'AQ = B.$$

Por lo tanto, queda demostrado el siguiente teorema:

Una forma cuadrática en n indeterminadas de matriz A, después de efectuar una transformación lineal de las indeterminadas de matriz Q, se convierte en una forma cuadrática en las nuevas indeterminadas, siendo la matriz de esta forma el producto $Q'AQ$.

Supongamos ahora que se efectúa una transformación lineal **no degenerada**, o sea, que la matriz Q y, por lo tanto, también la matriz Q', no son degeneradas. En este caso, se obtiene el producto $Q'AQ$ multiplicando la matriz A por unas matrices no degeneradas, por lo cual, como se deduce de los resultados del § 14, el rango de este producto es igual al rango de la matriz A. Por lo tanto, *al efectuar una transformación lineal no degenerada, el rango de la forma cuadrática no se altera.*

Veamos ahora, por analogía con el problema geométrico de la reducción de la ecuación de una curva central de segundo orden a la forma canónica (3), el problema de la reducción de una forma cuadrática arbitraria a la forma de una suma de cuadrados de las indeterminadas, o sea, a una forma en que todos los coeficientes de los productos de diversas indeterminadas sean iguales a cero, realizando para esto una transformación lineal no degenerada; esta forma especial de la forma cuadrática se llama *canónica*. Supongamos primero que, mediante una transformación lineal no degenerada, la forma cuadrática f en n indeterminadas $x_1, x_2, \ldots, x_n$ queda reducida a la forma canónica.

$$f = b_1y_1^2 + b_2y_2^2 + \ldots + b_ny_n^2, \tag{11}$$

donde $y_1, y_2, \ldots, y_n$ son las nuevas indeterminadas. Claro, algunos de los coeficientes $b_1, b_2, \ldots, b_n$ pueden ser iguales a cero. Demostremos que el *número de coeficientes en* (11), *diferentes de cero, es indispensablemente igual al rango* r *de la forma* f.

En efecto, como hemos llegado a la (11) mediante una transformación no degenerada, la forma cuadrática que figura en el segundo miembro de la igualdad (11) también tiene que ser de rango r. Sin embargo, la matriz de esta forma cuadrática tiene la forma diagonal

$$\begin{pmatrix} b_1 & & & 0 \\ & b_2 & & \\ & & \ddots & \\ 0 & & & b_n \end{pmatrix},$$

y la exigencia de que esta matriz tenga el rango r es equivalente a la suposición de que en su diagonal principal figuren exactamente r elementos diferentes de cero.

Pasemos a la demostración del siguiente **teorema fundamental sobre las formas cuadráticas.**

Toda forma cuadrática puede ser reducida a la forma canónica mediante una transformación lineal no degenerada. Si es que se considera una forma cuadrática real, todos los coeficientes de la transformación lineal indicada se pueden suponer reales.

Este teorema subsiste para el caso de formas cuadráticas en una indeterminada, puesto que son de la forma ax^2, que ya es canónica. Por consiguiente, podemos hacer la demostración por inducción sobre el número de indeterminadas, es decir, demostrar el teorema para las formas cuadráticas en n indeterminadas, suponiendo que ya está demostrado para las formas de un número menor de indeterminadas.

Sea dada una forma cuadrática

$$f=\sum_{i=1}^{n}\sum_{j=1}^{n}a_{ij}x_ix_j \qquad (12)$$

en n indeterminadas $x_1, x_2, \ldots, x_n$. Vamos a procurar hallar una transformación lineal no degenerada de tal modo que separe de f el cuadrado de una de las indeterminadas, o sea, que convierta a f en una suma de este cuadrado y una forma cuadrática en las demás indeterminadas. *Este objetivo se consigue fácilmente cuando entre* los coeficientes $a_{11}, a_{22}, \ldots, a_{nn}$ que figuran en la diagonal principal de la matriz de la forma f haya alguno diferente de cero, o sea, cuando en (12) haya por lo menos un cuadrado de una indeterminada x_i cuyo coeficiente sea diferente de cero.

Sea, por ejemplo, $a_{11}\neq 0$. Entonces, como fácilmente se comprueba, la expresión $a_{11}^{-1}(a_{11}x_1+a_{12}x_2+\ldots+a_{1n}x_n)^2$, que representa una forma cuadrática, contiene los mismos términos en la indeterminada x_1 que nuestra forma f y, por lo tanto, la diferencia que

$$f-a_{11}^{-1}(a_{11}x_1+a_{12}x_2+\ldots+a_{1n}x_n)^2=g$$

será una forma cuadrática que contendrá solamente a las indeterminadas $x_2, \ldots, x_n$, pero no a x_1. De aquí, que

$$f=a_{11}^{-1}(a_{11}x_1+a_{12}x_2+\ldots+a_{1n}x_n)^2+g.$$

Haciendo las notaciones

$$y_1=a_{11}x_1+a_{12}x_2+\ldots+a_{1n}x_n, \quad y_i=x_i \text{ para } i=2, 3, \ldots, n, \qquad (13)$$

se obtiene

$$f=a_{11}^{-1}y_1^2+g, \qquad (14)$$

donde g será ahora una forma cuadrática en las indeterminadas $y_2, y_3, \ldots, y_n$. La expresión (14) es la buscada para la forma f, *puesto que se ha obtenido de (12) mediante una transformación* lineal no degenerada, inversa a la transformación lineal (13), cuyo determinante es a_{11}, lo que implica que no sea degenerada.

Si se cumplen las igualdades $a_{11}=a_{22}=\ldots=a_{nn}=0$, se debe efectuar previamente una transformación lineal auxiliar que dé lugar a la aparición de cuadrados de las indeterminadas en nuestra forma f. Como entre los coeficientes de la expresión (12) tiene que haber diferentes de cero — en caso contrario no habría que demostrar nada — supondremos que, por ejemplo, $a_{12}\neq 0$, o sea, que f es la suma del término $2a_{12}x_1x_2$ y de otros términos, en cada uno de los cuales figura por lo menos una de las indeterminadas $x_3, \ldots, x_n$.

Hagamos ahora la transformación lineal

$$x_1=z_1-z_2, \; x_2=z_1+z_2, \; x_i=z_i \text{ para } i=3, \ldots, n. \qquad (15)$$

Esta no es degenerada, puesto que su determinante es:

$$\begin{vmatrix} 1 & -1 & 0 & \dots & 0 \\ 1 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 1 & \dots & 1 \end{vmatrix} = 2 \neq 0.$$

Como resultado de esta transformación, el término $2a_{12}x_1x_2$ tomará la forma

$$2a_{12}x_1x_2 = 2a_{12}(z_1 - z_2)(z_1 + z_2) = 2a_{12}z_1^2 - 2a_{12}z_2^2,$$

es decir, en la forma f aparecerán a la vez los cuadrados de dos indeterminadas con coeficientes diferentes de cero, que no podrán simplificarse con los demás términos, puesto que en cada uno de estos últimos hay por lo menos una de las indeterminadas $z_3, \dots, z_n$. Ahora nos encontramos en las condiciones del caso considerado anteriormente, de modo que con otra transformación lineal más, no degenerada, se podrá reducir la forma f a la forma (14).

Para terminar la demostración no queda más que señalar que la forma cuadrática g depende de un número menor que n de indeterminadas, y por la hipótesis de la inducción, se reduce a la forma canónica mediante una transformación no degenerada de las indeterminadas $y_2, y_3, \dots, y_n$. Esta transformación, considerada como una transformación (que, como fácilmente se comprueba, no es degenerada) de todas las n indeterminadas, según la cual y_1 se mantiene invariable, reduce (14) a la forma canónica. Por lo tanto, mediante dos o tres transformaciones lineales no degeneradas (que se pueden sustituir por una sola transformación no degenerada: por su producto), la forma cuadrática f se reduce a una suma de cuadrados de las indeterminadas con ciertos coeficientes. Como ya sabemos, el número de estos cuadrados es igual al rango r de la forma. Si además de esto, la forma cuadrática f es real, los coeficientes en la forma canónica de f, así como en la transformación lineal que reduce f a esta forma, serán reales; en efecto, tanto la transformación lineal inversa de (13) como la transformación lineal (15) tienen coeficientes reales.

El teorema queda demostrado fundamental. El método utilizado en esta demostración puede aplicarse en ejemplos concretos para la reducción efectiva de una forma cuadrática a la forma canónica. Pero, en lugar de la inducción que se empleaba en la demostración, se aplica el método expuesto para separar sucesivamente los cuadrados de las indeterminadas.

Ejemplo. Reducir la forma cuadrática

$$f = 2x_1x_2 - 6x_2x_3 + 2x_3x_1 \qquad (16)$$

a la forma canónica.

Debido a la ausencia en esta forma de los cuadrados de las indeterminadas, efectuamos primero la transformación lineal no degenerada

$$x_1=y_1-y_2,\quad x_1=y_1+y_2,\quad x_3=y_3$$

de matriz

$$A=\begin{pmatrix}1 & -1 & 0\\ 1 & 1 & 0\\ 0 & 0 & 1\end{pmatrix},$$

después de lo cual se obtiene:

$$f=2y_1^2-2y_2^2-4y_1y_3-8y_2y_3.$$

Ahora, el coeficiente de y_1^2 es diferente de cero y, por esto, en nuestra forma se puede separar el cuadrado de una indeterminada. Haciendo

$$z_1=2y_1-2y_3,\quad z_2=y_2,\quad z_3=y_3,$$

o sea, efectuando la transformación lineal cuya inversa tiene la matriz

$$B=\begin{pmatrix}\frac{1}{2} & 0 & 1\\ 0 & 1 & 0\\ 0 & 0 & 1\end{pmatrix},$$

la forma f se reduce a la forma

$$f=\frac{1}{2}z_1^2-2z_2^2-2z_3^2-8z_2z_3.$$

Por ahora solamente se ha separado el cuadrado de la indeterminada z_1, puesto que la forma contiene todavía el producto de las otras dos indeterminadas. Aplicando la desigualdad de cero del coeficiente de z_2^2, empleamos de nuevo el método expuesto anteriormente. Efectuando la transformación lineal

$$t_1=z_1,\quad t_2=-2z_2-4z_3,\quad t_3=z_3,$$

cuya inversa tiene la matriz

$$C=\begin{pmatrix}1 & 0 & 0\\ 0 & -\frac{1}{2} & -2\\ 0 & 0 & 1\end{pmatrix},$$

se reduce, finalmente, *la forma f a la forma canónica*

$$f=\frac{1}{2}t_1^2-\frac{1}{2}t_2^2+6t_3^2. \tag{17}$$

La transformación lineal que reduce simultáneamente la forma (16) a la forma (17) tiene por matriz el producto

$$ABC=\begin{pmatrix}\frac{1}{2} & \frac{1}{2} & 3\\ \frac{1}{2} & -\frac{1}{2} & -1\\ 0 & 0 & 1\end{pmatrix}.$$

Mediante una sustitución directa se puede comprobar que la transformación lineal no degenerada $\left(\text{puesto que el determinante es igual a } -\frac{1}{2}\right)$

$$x_1 = \frac{1}{2} t_1 + \frac{1}{2} t_2 + 3t_3,$$
$$x_2 = \frac{1}{2} t_1 - \frac{1}{2} t_2 - t_3,$$
$$x_3 = \qquad\qquad t_3$$

transforma (16) en (17).

La teoría de la reducción de una forma cuadrática a la forma canónica se ha elaborado por analogía con la teoría geométrica de las curvas centrales de segundo orden, pero no puede suponerse que es una generalización de esta última. En efecto, en nuestra teoría se permitía la aplicación de cualesquiera transformaciones lineales no degeneradas, mientras que la reducción de la ecuación de una curva de segundo orden a la forma canónica se consigue aplicando transformaciones lineales de una forma (2) muy especial, que representan rotaciones del plano. Sin embargo, esta teoría geométrica se puede generalizar al caso de formas cuadráticas en n indeterminadas con coeficientes **reales.** En el cap. 8 se hará una exposición de esta generalización, denominada **reducción de las formas cuadráticas a los ejes principales.**

§ 27. Ley de inercia

Generalmente, la forma canónica a que se reduce una forma cuadrática dada no se determina unívocamente, pues toda forma cuadrática se puede reducir a la forma canónica de muchos modos. Así, la forma cuadrática $f = 2x_1x_2 - 6x_2x_3 + 2x_3x_1$, considerada en el párrafo anterior, mediante la transformación lineal no degenerada

$$x_1 = t_1 + 3t_2 + 2t_3,$$
$$x_2 = t_1 - t_2 - 2t_3,$$
$$x_3 = \qquad t_2$$

se reduce a la forma canónica

$$f = 2t_1^2 + 6t_2^2 - 8t_3^2,$$

distinta de la obtenida anteriormente.

Surge la pregunta: ¿Qué tienen de común las diversas formas cuadráticas canónicas a que se reduce la forma f dada? Como veremos, esta cuestión está estrechamente ligada con la siguiente pregunta: ¿cuál es la condición para que una de las dos formas cuadráticas dadas se reduzca a la otra mediante una transformación lineal?

Sin embargo, la respuesta a estas preguntas depende de que sean reales o complejas las formas cuadráticas consideradas.

Supongamos primero que se consideran formas cuadráticas *complejas* arbitrarias y que, a la vez, se permite el empleo de transformaciones lineales no degeneradas también con coeficientes complejos arbitrarios. *Ya sabemos que toda forma cuadrática f en n* indeterminadas de rango r, se reduce a la forma canónica

$$f = c_1y_1^2 + c_2y_2^2 + \ldots + c_ry_r^2,$$

donde todos los coeficientes $c_1, c_2, \ldots, c_r$ son diferentes de cero. Teniendo en cuenta que se puede extraer la raíz cuadrada de cualquier número complejo, realicemos la siguiente transformación lineal no degenerada:

$$z_i = \sqrt{c_i}\, y_i \text{ para } i = 1, 2, \ldots, r;\ z_j = y_j \text{ para } j = r+1, \ldots, n.$$

Esta reduce f a la forma

$$f = z_1^2 + z_2^2 + \ldots + z_r^2, \qquad (1)$$

denominada *normal*, la cual es, simplemente, la suma de los cuadrados de r indeterminadas con coeficientes iguales a la unidad.

La forma normal depende solamente del rango r de la forma f, es decir, todas las formas cuadráticas de rango r se reducen a una misma forma normal (1). Por consiguiente, si las formas f y g en n indeterminadas son de igual rango r, se puede reducir f a la forma (1), *y después, (1) a la forma g, lo que significa que existe una transformación* lineal no degenerada que reduce f a la forma g. Por otra parte, como una transformación lineal no degenerada nunca altera el rango de la forma, llegamos al resultado siguiente:

Dos formas cuadráticas complejas en n indeterminadas se reducen una a otra mediante transformaciones lineales no degeneradas con coeficientes complejos cuando, y sólo cuando, éstas son de un mismo rango.

De este teorema se deduce sin dificultad que *puede ser forma canónica de una forma cuadrática compleja de rango r cualquier suma de cuadrados de r indeterminadas con cualesquiera coeficientes complejos diferentes de cero.*

El asunto se complica *si se consideran formas cuadráticas* **reales** y, sobre todo, si se permiten solamente transformaciones lineales con coeficientes reales, cosa muy importante. En este caso, ya no se puede reducir cualquier forma a la forma (1), puesto que probablemente se tendría que efectuar la extracción de la raíz cuadrada de un número negativo. Sin embargo, si llamamos ahora **forma normal** de una forma cuadrática a la suma de los cuadrados de unas cuantas *indeterminadas, tomadas con los coeficientes $+1$ o -1, se puede* demostrar fácilmente, que *cualquier forma cuadrática real f se puede*

reducir a la forma normal mediante una transformación lineal no degenerada con coeficientes reales.

En efecto, la forma f en n indeterminadas, de rango r, se reduce a la forma canónica, que se puede escribir del modo siguiente (cambiando la numeración de las indeterminadas, si fuese necesario):

$$f = c_1y_1^2 + \ldots + c_ky_k^2 - c_{k+1}y_{k+1}^2 - \ldots - c_ry_r^2, \qquad 0 \leqslant k \leqslant r,$$

donde todos los números $c_1, \ldots, c_k; c_{k+1}, \ldots, c_r$, son diferentes de cero y positivos. Entonces, la transformación lineal no degenerada con coeficientes reales $z_i = \sqrt{c_i}y_i$ para $i = 1, 2, \ldots, r$, $z_j = y_j$ para $j = r + 1, \ldots, n$, reduce f a la forma normal,

$$f = z_1^2 + \ldots + z_k^2 - z_{k+1}^2 - \ldots - z_r^2.$$

El número total de cuadrados que figuran aquí es igual al rango de la forma.

Una forma cuadrática real se puede reducir a la forma normal mediante muchas transformaciones diversas, pero salvo el orden de numeración de las indeterminadas ésta se reduce solamente a una forma normal. Esto lo muestra el importante teorema, denominado **ley de inercia de las formas cuadráticas reales:**

El número de cuadrados positivos, así como el número de cuadrados negativos, en la forma normal a que se reduce una forma cuadrática dada con coeficientes reales por una transformación lineal real no degenerada, no depende de la elección de esta transformación.

En efecto, supongamos que una forma cuadrática f de rango r, en n indeterminadas $x_1, x_2, \ldots, x_n$ se ha reducido a la forma normal de dos modos diversos:

$$f = y_1^2 + \ldots + y_k^2 - y_{k+1}^2 - \ldots - y_r^2 =$$
$$= z_1^2 + \ldots + z_l^2 - z_{l+1}^2 - \ldots - z_r^2. \qquad (2)$$

Como el paso de las indeterminadas $x_1, x_2, \ldots, x_n$ a las indeterminadas $y_1, y_2, \ldots, y_n$ era una transformación lineal no degenerada, las segundas indeterminadas también se expresarán linealmente mediante las primeras con un determinante diferente de cero:

$$y_i = \sum_{s=1}^{n} a_{is}x_s, \qquad i = 1, 2, \ldots, n. \qquad (3)$$

Análogamente

$$z_j = \sum_{t=1}^{n} b_{jt}x_t, \qquad j = 1, 2, \ldots, n, \qquad (4)$$

donde el determinante de los coeficientes es de nuevo diferente de cero. Los coeficientes en (3), al igual que en (4), son números reales.

Supongamos ahora que $k < l$, y escribamos el sistema de igualdades:

$$y_1 = 0, \ldots, y_k = 0, \quad z_{l+1} = 0, \ldots, z_r = 0, \ldots, z_n = 0. \qquad (5)$$

Si los primeros miembros de estas igualdades se sustituyen por sus expresiones (3) y (4), se obtiene un sistema de $n - l + k$ ecuaciones lineales homogéneas con n incógnitas $x_1, x_2, \ldots, x_n$. El número de ecuaciones en este sistema es menor que el número de incógnitas, por consiguiente, por el § 1, este sistema posee solución real *no nula*, $\alpha_1, \alpha_2, \ldots, \alpha_n$.

Sustituyamos ahora en la igualdad (2) todas las y y todas las z por sus expresiones (3) y (4), y pongamos después en lugar de las indeterminadas los números $\alpha_1, \alpha_2, \ldots, \alpha_n$. Si, para abreviar, se designan con $y_i(\alpha)$ y $z_j(\alpha)$ los valores de las indeterminadas y_i y z_j, que se obtienen después de esta sustitución, la igualdad (2) se convierte, en virtud de (5), en la igualdad

$$-y_{k+1}^2(\alpha) - \ldots - y_r^2(\alpha) = z_1^2(\alpha) + \ldots + z_l^2(\alpha). \qquad (6)$$

Como todos los coeficientes en (3) y (4) son reales, todos los cuadrados que figuran en la igualdad (6) son positivos; por consiguiente, de la igualdad (6) se deduce la igualdad a cero de todos estos cuadrados; de aquí resultan las igualdades:

$$z_1(\alpha) = 0, \ldots, z_l(\alpha) = 0. \qquad (7)$$

Por otra parte, debido a la misma elección de los números $\alpha_1, \alpha_2, \ldots, \alpha_n$

$$z_{l+1}(\alpha) = 0, \ldots, z_r(\alpha) = 0, \ldots, z_n(\alpha) = 0. \qquad (8)$$

Por lo tanto, en virtud de (7) y (8), el sistema de n ecuaciones lineales homogéneas

$$z_i = 0, \quad i = 1, 2, \ldots, n,$$

con n incógnitas $x_1, x_2, \ldots, x_n$ posee solución no nula $\alpha_1, \alpha_2, \ldots$ $\ldots, \alpha_n$, por lo cual, el determinante de este sistema tiene que ser igual a cero. Sin embargo, esto es absurdo, puesto que se suponía que la transformación (4) era no degenerada. Suponiendo que $l < k$ llegamos también al absurdo. De aquí se deduce la igualdad $k = l$, que demuestra el teorema.

El número de cuadrados positivos en la forma normal a que se reduce una forma cuadrática real f dada, se llama *índice positivo de inercia* de esta forma; el número de cuadrados negativos, *índice negativo de inercia*; y la diferencia entre el índice positivo y el negativo, *signatura* de la forma f. Está claro que dado el rango de la forma, cualquiera de los tres números que acabamos de definir determina completamente a los otros dos y, por esto, en los enunciados que siguen se puede mencionar cualquiera de ellos.

Demostremos ahora el siguiente **teorema:**

Dos formas cuadráticas en n indeterminadas con coeficientes reales se reducen una a otra mediante transformaciones lineales reales no degeneradas cuando, y sólo cuando, tienen el mismo rango y la misma signatura.

En efecto, supongamos que la forma f se reduce a la forma g mediante una transformación real no degenerada. Ya se sabe que esta transformación no altera el rango de la forma. Esta tampoco puede alterar la signatura, puesto que, en caso contrario, las formas f y g se reducirían a diferentes formas normales, y la forma f se reduciría a ambas formas normales, lo cual contradice a la ley de inercia. Recíprocamente, si las formas f y g tienen un mismo rango y una misma signatura, entonces se reducen a una misma forma normal y, en consecuencia, se pueden reducir una a otra.

Si se da una forma cuadrática g en la forma canónica

$$g = b_1y_1^2 + b_2y_2^2 + \ldots + b_ry_r^2, \tag{9}$$

con coeficientes reales diferentes de cero, el rango es, evidentemente, igual a r. Fácilmente se ve, empleando el método aplicado anteriormente de reducción de esa forma a la forma normal, que el índice positivo de inercia de la forma g es igual al número de coeficientes positivos en el segundo miembro de la igualdad (9). De aquí y del teorema anterior se deduce el siguiente resultado:

La forma (9) *será la forma canónica de una forma cuadrática f dada cuando, y sólo cuando, el rango de ésta sea igual a r y su índice positivo de inercia coincida con el número de coeficientes positivos en* (9).

Formas cuadráticas descomponibles. Multiplicando dos formas lineales cualesquiera en n indeterminadas,

$$\varphi = a_1x_1 + a_2x_2 + \ldots + a_nx_n, \quad \psi = b_1x_1 + b_2x_2 + \ldots + b_nx_n,$$

se obtiene, evidentemente, una forma cuadrática. No cualquier forma cuadrática se puede representar en forma de un producto de dos formas lineales, y queremos deducir las condiciones para que esto tenga lugar, o sea, para que la forma cuadrática sea *descomponible*.

Una forma cuadrática compleja $f(x_1, x_2, \ldots, x_n)$ es descomponible cuando, y sólo cuando, su rango es menor o igual a dos. Una forma cuadrática real $f(x_1, x_2, \ldots, x_n)$ es descomponible cuando, y sólo cuando, su rango no es mayor que la unidad, o es igual a dos, pero su signatura es igual a cero.

Examinemos primero el producto de las formas lineales φ y ψ. Si al menos una de estas formas es nula, su producto también será una forma cuadrática con coeficientes nulos, es decir, tendrá el rango 0. Si las formas lineales φ y ψ son proporcionales,

$$\psi = c\varphi,$$

siendo $c \neq 0$, y la forma φ no es nula, supondremos que a_1, por ejemplo, es diferente de cero. Entonces la transformación lineal no degenerada

$$y_1 = a_1x_1 + \ldots + a_nx_n, \quad y_i = x_i \text{ para } i = 2, 3, \ldots, n$$

reduce la forma cuadrática $\varphi\psi$ a la forma

$$\varphi\psi = cy_1^2.$$

Como en el segundo miembro figura una forma cuadrática de rango 1, la forma cuadrática $\varphi\psi$ también tendrá el rango 1. Finalmente, si las formas lineales φ y ψ no son proporcionales, supondremos que

$$\begin{vmatrix} a_1 & a_2 \\ b_1 & b_2 \end{vmatrix} \neq 0.$$

Entonces la transformación lineal

$$\begin{aligned} y_1 &= a_1x_1 + a_2x_2 + \ldots + a_nx_n, \\ y_2 &= b_1x_1 + b_2x_2 + \ldots + b_nx_n, \\ y_i &= x_i \text{ para } i = 3, 4, \ldots, n \end{aligned}$$

no será degenerada; ésta reducirá la forma cuadrática $\varphi\psi$ a la forma

$$\varphi\psi = y_1y_2.$$

En el segundo miembro figura una forma cuadrática de rango 2, que en el caso de coeficientes reales tendrá la signatura 0.

Pasemos a la demostración de la afirmación recíproca. Por supuesto, una forma cuadrática de rango 0 puede considerarse como el producto de dos formas lineales, una de las cuales es nula. Luego, una forma cuadrática $f(x_1, x_2, \ldots, x_n)$ de rango 1, mediante una transformación lineal no degenerada, se reduce a la forma

$$f = cy_1^2, \quad c \neq 0,$$

o sea, a la forma

$$f = (cy_1)\, y_1.$$

Expresando linealmente y_1 mediante $x_1, x_2, \ldots, x_n$, se obtiene la representación de la forma f en un producto de dos formas lineales. Finalmente, una forma cuadrática real $f(x_1, x_2, \ldots, x_n)$ de rango 2 y signatura 0, mediante una transformación lineal no degenerada, se reduce a la forma

$$f = y_1^2 - y_2^2;$$

a esta misma forma se puede reducir cualquier forma cuadrática compleja de rango 2. Sin embargo,

$$y_1^2 - y_2^2 = (y_1 - y_2)(y_1 + y_2),$$

y en el segundo miembro, después de sustituir y_1 y y_2 por sus expresiones lineales mediante $x_1, x_2, \ldots, x_n$, resultará el producto de dos formas lineales. Así, el teorema queda demostrado.

§ 28. Formas definidas positivas

Una forma cuadrática f en n indeterminadas con coeficientes reales se llama *definida positiva*, si se reduce a una forma normal que consta de n cuadrados positivos, es decir, si tanto el rango como el índice positivo de inercia de esta forma son iguales al número de las indeterminadas.

El siguiente teorema da la posibilidad de caracterizar las formas definidas positivas, sin reducirlas a la forma normal o canónica.

Una forma cuadrática f en n indeterminadas $x_1, x_2, \ldots, x_n$ con coeficientes reales, es definida positiva si, y sólo si, para cualesquiera valores reales de estas indeterminadas, no simultáneamente nulos, la forma toma valores positivos.

Demostración. Supongamos que la forma f es definida positiva, o sea, que se reduce a la forma normal

$$f = y_1^2 + y_2^2 + \ldots + y_n^2, \tag{1}$$

donde

$$y_i = \sum_{j=1}^{n} a_{ij}x_j, \qquad i = 1, 2, \ldots, n, \tag{2}$$

siendo diferente de cero el determinante de los coeficientes reales a_{ij}. Si se quieren poner en f valores reales arbitrarios de las indeterminadas $x_1, x_2, \ldots, x_n$, al menos uno de los cuales es diferente de cero, se pueden ponerlos primero en (2) y, después, los valores obtenidos de y_i, en (1). Obsérvese que los valores obtenidos en (2) para $y_1, y_2, \ldots, y_n$, no pueden ser simultáneamente iguales a cero, puesto que en caso contrario resultaría que el sistema de ecuaciones lineales homogéneas

$$\sum_{j=1}^{n} a_{ij}x_j = 0, \qquad i = 1, 2, \ldots, x,$$

poseería solución no nula a pesar de que su determinante es diferente de cero. Sustituyendo en (1) los valores obtenidos de $y_1, y_2, \ldots, y_n$ se obtiene un valor de la forma f, igual a la suma de los cuadrados de n números reales que no son todos iguales a cero; por consiguiente, este valor es estrictamente positivo.

Recíprocamente, supongamos que la forma f no es definida positiva, o sea, que su rango o su índice positivo de inercia es menor que n. Esto significa que en su forma normal, a la que se reduce mediante una transformación lineal no degenerada (2), el cuadrado de al menos una de las indeterminadas, por ejemplo, de y_n,

o bien falta, o bien figura con el signo menos. Demostremos que en este caso se pueden elegir para las indeterminadas $x_1, x_2, \ldots, x_n$ unos valores reales, no todos iguales a cero, de modo que el valor de esta forma para estos valores de las indeterminadas sea igual a cero e incluso negativo. Tales son, por ejemplo, los valores que se obtienen para $x_1, x_2, \ldots, x_n$ al resolver por la regla de Cramer el sistema de ecuaciones lineales que resulta de (2) para $y_1 = y_2 = \ldots = y_{n-1} = 0$, $y_n = 1$. En efecto, para estos valores de las indeterminadas $x_1, x_2, \ldots, x_n$ la forma es igual a cero, si y_n^2 no figura en la forma normal, e igual a -1, si y_n^2 figura en la forma normal con el signo menos.

El teorema que acabamos de demostrar se emplea en todos los casos donde se aplican las formas cuadráticas definidas positivas. Sin embargo, con su ayuda no se puede determinar, valiéndose de los coeficientes de la forma, si ésta es definida positiva o no. Para este fin sirve otro teorema que enunciaremos y demostraremos después de que se introduzca un concepto auxiliar.

Sea dada una forma cuadrática f en n indeterminadas de matriz $A = (a_{ij})$. Los menores de orden $1, 2, \ldots, n$ de esta matriz, situados en el ángulo superior de la izquierda, o sea, los menores

$$a_{11}, \quad \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}, \ldots, \begin{vmatrix} a_{11} & a_{12} & \ldots & a_{1k} \\ a_{21} & a_{22} & \ldots & a_{2k} \\ \cdot & \cdot & \cdot & \cdot \\ a_{k1} & a_{k2} & \ldots & a_{kk} \end{vmatrix}, \ldots, \begin{vmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \cdot & \cdot & \cdot & \cdot \\ a_{n1} & a_{n2} & \ldots & a_{nn} \end{vmatrix},$$

el último de los cuales, evidentemente, coincide con el determinante de la matriz A, se llaman *menores principales* de la forma f.

Subsiste el siguiente **teorema:**

Una forma cuadrática f en n indeterminadas con coeficientes reales es definida positiva si, y sólo si, todos sus menores principales son estrictamente positivos.

Demostración. El teorema es cierto para $n = 1$, puesto que en este caso la forma es ax^2 y, por lo tanto, es definida positiva si, y sólo si, $a > 0$. Por esta razón demostraremos el teorema para el caso de n indeterminadas, suponiendo que ya está demostrado para las formas cuadráticas en $n - 1$ indeterminadas.

Hagamos primero la observación siguiente:

Si una forma cuadrática f con coeficientes reales que forman una matriz A, se somete a una transformación lineal no degenerada de matriz real Q, *el signo del determinante de la forma (o sea, del determinante de su matriz) no varía.*

En efecto, después de la transformación se obtiene una forma cuadrática cuya matriz es $Q'AQ$; pero, como $|Q'| = |Q|$, resulta:

$$|Q'AQ| = |Q'| \cdot |A| \cdot |Q| = |A| \cdot |Q|^2,$$

o sea, el determinante $|A|$ se multiplica por un número positivo.

Supongamos ahora que se ha dado una forma cuadrática

$$f=\sum_{i,\,j=1}^{n} a_{ij}x_ix_j.$$

Esta se puede escribir en la forma

$$f=\varphi(x_1, x_2, \ldots, x_{n-1})+2\sum_{i=1}^{n-1} a_{in}x_ix_n+a_{nn}x_n^2, \qquad (3)$$

donde φ es una forma cuadrática en $n-1$ indeterminadas, formada por los términos de la forma f que no contienen a la indeterminada x_n. Obsérvese que los menores principales de la forma φ coinciden con todos los menores principales de la forma f, menos con el último.

Supongamos que la forma f es definida positiva. En este caso, la forma φ también será definida positiva, pues si existiesen unos valores de las indeterminadas $x_1, x_2, \ldots, x_n$, no simultáneamente nulos, para los que la forma φ tomase un valor no estrictamente positivo, entonces, poniendo complementariamente $x_n=0$, en virtud de (3), se obtendría también un valor no estrictamente positivo para la forma f, a pesar de que no todos los valores de las indeterminadas $x_1, x_2, \ldots, x_{n-1}, x_n$ son iguales a cero. En consecuencia, según la hipótesis de inducción, todos los menores principales de la forma φ, es decir, todos los menores principales de la forma f, menos el último, son estrictamente positivos. En lo que se refiere al último menor principal de la forma f, o sea, al determinante de la misma matriz A, éste es positivo debido a las razones siguientes: como la forma f es definida positiva, mediante una transformación lineal no degenerada, ésta se reduce a la forma normal que consta de n cuadrados positivos. El determinante de esta forma normal es estrictamente positivo y, por esto, en virtud de la observación hecha anteriormente, es también positivo el determinante de la misma forma f.

Supongamos ahora que todos los menores principales de la forma f son estrictamente positivos. De aquí se deduce que son positivos todos los menores principales de la forma φ, y por la hipótesis de inducción, ésta es definida positiva. Por consiguiente, existe una transformación lineal no degenerada de las indeterminadas $x_1, x_2, \ldots, x_{n-1}$ que reduce la forma φ a una suma de $n-1$ cuadrados positivos de las nuevas indeterminadas $y_1, y_2, \ldots, y_{n-1}$. Esta transformación lineal se puede completar hasta una transformación lineal (no degenerada) de todas las indeterminadas $x_1, x_2, \ldots, x_n$, haciendo $x_n=y_n$. En virtud de (3), mediante la transformación indicada, la forma f se reduce a la forma

$$f=\sum_{i=1}^{n-1} y_i^2+2\sum_{i=1}^{n-1} b_{in}y_iy_n+b_{nn}y_n^2; \qquad (4)$$

Las expresiones exactas de los coeficientes b_{in} no tienen interés alguno. Como

$$y_i^2 + 2b_{in}y_iy_n = (y_i + b_{in}y_n)^2 - b_{in}^2y_n^2,$$

en virtud de (4), la transformación lineal no degenerada

$$z_i = y_i + b_{in}y_n, \qquad i = 1, 2, \ldots, n-1,$$
$$z_n = \qquad y_n$$

reduce la forma f a la forma canónica

$$f = \sum_{i=1}^{n-1} z_i^2 + cz_n^2. \qquad (5)$$

Para demostrar que la forma f es definida positiva, no queda más que demostrar que el número c es positivo. El determinante de la forma que figura en el segundo miembro de la igualdad (5) es igual a c. Sin embargo, este determinante tiene que ser positivo, puesto que el segundo miembro de la igualdad (5) se ha obtenido de la forma f mediante dos transformaciones lineales no degeneradas, y el determinante de la forma f, como último de los menores principales de ésta, es positivo.

Así, pues, el teorema queda demostrado.

Ejempos. 1. La forma cuadrática

$$f = 5x_1^2 + x_2^2 + 5x_3^2 + 4x_1x_2 - 8x_1x_3 - 4x_2x_3$$

es definida positiva, ya que sus menores principales

$$5, \quad \begin{vmatrix} 5 & 2 \\ 2 & 1 \end{vmatrix} = 1, \quad \begin{vmatrix} 5 & 2 & -4 \\ 2 & 1 & -2 \\ -4 & -2 & 5 \end{vmatrix} = 1$$

son positivos.

2. La forma cuadrática

$$f = 3x_1^2 + x_2^2 + 5x_3^2 + 4x_1x_2 - 8x_1x_3 - 4x_2x_3$$

no es definida positiva, puesto que su segundo menor principal es negativo

$$\begin{vmatrix} 3 & 2 \\ 2 & 1 \end{vmatrix} = -1.$$

Obsérvese que por analogía con las formas cuadráticas definidas positivas se pueden definir las *formas definidas negativas*, o sea, unas formas cuadráticas no degeneradas con coeficientes reales cuyas formas normales contienen solamente cuadrados negativos de las indeterminadas. Las formas cuadráticas degeneradas, cuyas formas normales constan de cuadrados de un mismo signo, se llaman a veces *semidefinidas*. Finalmente, las formas cuadráticas, cuyas formas normales contienen cuadrados de las indeterminadas, tanto positivos como negativos, son *indefinidas*.

CAPITULO VII

ESPACIOS LINEALES

§ 29. Definición del espacio lineal. Isomorfismo

La definición de espacio vectorial de n dimensiones dada en el § 8, comenzaba con la definición de un vector de n dimensiones como un sistema ordenado de n números. Para los vectores de n dimensiones se definieron luego la suma y el producto de ellos por números, lo cual condujo a la noción de espacio vectorial de n dimensiones. Los primeros ejemplos de espacios vectoriales son los conjuntos de vectores-segmentos que parten del origen de coordenadas, en el plano o en el espacio tridimensional. Sin embargo, tratando estos ejemplos en el curso de geometría, no siempre creemos necesario determinar los vectores por sus componentes respecto a un sistema fijo de coordenadas, puesto que la suma de vectores y su producto por un escalar se determinan geométricamente, independientemente de la elección del sistema de coordenadas. Precisamente, la suma de vectores en el plano o en el espacio se efectúa según la regla del paralelogramo y el producto de un vector por un número α significa el alargamiento de este vector en α veces (o la contracción, teniendo que cambiar la dirección del vector por la contraria en caso de que α sea negativo). En el caso general, también es conveniente hacer una definición «sin recurrir a coordenadas» del espacio vectorial, es decir, hacer una definición que no necesite determinar los vectores como sistemas ordenados de números. Ahora se dará tal definición. Esta definición es **axiomática**, pues en ella no se tratará de las propiedades de cada vector por separado, sino que se enumerarán las propiedades que deben poseer las operaciones con los vectores.

Sea dado un conjunto V; sus elementos se designarán con letras latinas minúsculas: $a, b, c, \ldots$*. Supongamos también que en el conjunto V se han definido las operaciones siguientes: *la suma*, que pone en correspondencia a cada par de elementos a, b de V un elemento unívocamente determinado $a + b$ de V, denominado *suma*,

* A diferencia de lo convenido en el capítulo 2, en el presente capítulo y en el siguiente, los vectores se designarán con letras latinas minúsculas, mientras que los números, con letras griegas minúsculas.

y la *multiplicación de un elemento por un número real*, según la cual, el *producto* αa *del elemento* a por el número α está unívocamente determinado y pertenece a V.

Los elementos del conjunto V se llamarán *vectores* y el mismo conjunto V, *espacio lineal* (o *vectorial*, o *afín*) *real*, si las operaciones indicadas poseen las propiedades I—VIII que siguen:

I. La suma es conmutativa, $a + b = b + a$.

II. La suma es asociativa, $(a + b) + c = a + (b + c)$.

III. Existe en V un *elemento nulo* (cero) 0, que satisface a la condición: $a + 0 = a$ para todos los a de V.

Aplicando I, es fácil demostrar la *unicidad del elemento nulo*: si 0_1 y 0_2 son dos elementos nulos, se tiene:

$$0_1 + 0_2 = 0_1, \qquad 0_1 + 0_2 = 0_2 + 0_1 = 0_2,$$

de donde $0_1 = 0_2$.

IV. Para todo elemento a de V existe un *elemento opuesto* $-a$, que satisface a la condición: $a + (-a) = 0$.

Fácilmente se comprueba, en virtud de II y I, la *unicidad del elemento opuesto*: si $(-a)_1$ y $(-a)_2$ son dos elementos opuestos de a, entonces,

$$(-a)_1 + [a + (-a)_2] = (-a_1) + 0 = (-a)_1,$$
$$[(-a)_1 + a] + (-a)_2 = 0 + (-a)_2 = (-a)_2,$$

de donde $(-a)_1 = (-a)_2$.

De los axiomas I—IV se deduce la *existencia y unicidad de la diferencia* $a - b$, o sea, de un elemento que satisface a la ecuación

$$b + x = a. \tag{1}$$

En efecto, se puede poner

$$a - b = a + (-b),$$

pues

$$b + [a + (-b)] = [b + (-b)] + a = 0 + a = a.$$

Si existiese otro elemento más, c, que satisfaciese a la ecuación (1), o sea, que

$$b + c = a,$$

agregando a los dos miembros de esta igualdad el elemento $-b$, resultaría

$$c = a + (-b).$$

Los axiomas V—VIII que siguen (compárese con el § 8), ligan la multiplicación por un número con la suma y con las operaciones sobre los números. Precisamente, para cualesquiera elementos a, b de V, para cualesquiera números reales α, β y para el número real 1, *se tienen que cumplir las igualdades:*

V. $$\alpha (a + b) = \alpha a + \alpha b;$$

VI. $$(\alpha+\beta)\,a=\alpha a+\beta a;$$

VII. $$(\alpha\beta)\,a=\alpha\,(\beta a);$$

VIII. $$1\cdot a=a.$$

Señalemos algunas de las propiedades elementales de estos axiomas:

[1] $$\alpha\cdot 0=0.$$

En efecto, para un a de V,

$$\alpha a=\alpha(a+0)=\alpha a+\alpha\cdot 0,$$

o sea,

$$\alpha\cdot 0=\alpha a-\alpha a=\alpha a+[-(\alpha a)]=0,$$

[2] $$0\cdot a=0,$$

donde en el primer miembro figura el número cero, mientras que en el segundo miembro, el elemento nulo de V.

Para la demostración, tomemos cualquier número α. Entonces

$$\alpha a=(\alpha+0)\,a=\alpha a+0\cdot a,$$

de donde

$$0\cdot a=\alpha a-\alpha a=0.$$

[3]. Si $\alpha a=0$, entonces $\alpha=0$, o bien $a=0$. En efecto, si $\alpha\neq 0$, es decir, que existe el número α^{-1}, entonces

$$a=1\cdot a=(\alpha^{-1}\alpha)\,a=\alpha^{-1}(\alpha a)=\alpha^{-1}\cdot 0=0.$$

[4]. $$\alpha\,(-a)=-\alpha a.$$

En efecto,

$$\alpha a+\alpha\,(-a)=\alpha\,[a+(-\alpha)]=\alpha\cdot 0=0,$$

o sea, el elemento $\alpha\,(-a)$ es opuesto al elemento αa.

[5]. $$(-\alpha)\,a=-\alpha a.$$

En realidad,

$$\alpha a+(-\alpha)\,a=[\alpha+(-a)]\,a=0\cdot a=0,$$

o sea, el elemento $(-\alpha)\,a$ es opuesto al elemento αa.

[6]. $$\alpha\,(a-b)=\alpha a-\alpha b.$$

En efecto, en virtud de [4], tendremos

$$\alpha\,(a-b)=\alpha\,[a+(-b)]=\alpha a+\alpha\,(-b)=\alpha a+(-\alpha b)=\alpha a-\alpha b.$$

[7]. $$(\alpha-\beta)\,a=\alpha a-\beta a.$$

Se tiene, en efecto

$$(\alpha-\beta)\,a=[\alpha+(-\beta)]\,a=\alpha a+(-\beta)\,a=\alpha a+(-\beta a)=\alpha a-\beta a.$$

Obsérvese que los axiomas y las consecuencias enumeradas se emplearán a continuación sin reservas especiales.

Antes se dio la definición de espacio lineal real. Suponiendo que en el conjunto V no sólo se ha determinado el producto por números reales, sino también por cualesquiera números complejos, conservando los mismos axiomas I-VIII, se obtiene la definición de *espacio lineal complejo*. Para fijar ideas, se **examinarán a continuación los espacios lineales reales; sin emabargo, todo lo que se diga en el presente capítulo se refiere también palabra por palabra al caso de espacios lineales complejos.**

Es fácil señalar **ejemplos** de espacios lineales reales. Estos son, ante todo, los espacios vectoriales reales de n dimensiones formados por los vectores-filas, que se estudiaron en el cap. 2. También son espacios lineales los conjuntos de vectores-segmentos que parten del origen de coordenadas en el plano o en el espacio tridimensional, si las operaciones de suma y de multiplicación por un número se entienden en el sentido geométrico que se indicó al comienzo de este párrafo.

También existen ejemplos de espacios lineales, como suele decirse, de «infinitas dimensiones». Consideremos todas las sucesiones posibles de números reales; éstas son de la forma

$$a=(\alpha_1\alpha_2, \ldots, \alpha_n, \ldots).$$

Las operaciones con las sucesiones se efectúan componente a componente: si

$$b=(\beta_1, \beta_2, \ldots, \beta_n, \ldots),$$

se tiene

$$a+b=(\alpha_1+\beta_1, \alpha_2+\beta_2, \ldots, \alpha_n+\beta_n, \ldots);$$

por otra parte, para cualquier número real γ,

$$\gamma a=(\gamma\alpha_1, \gamma\alpha_2, \ldots, \gamma\alpha_n, \ldots).$$

Todos los axiomas I—VIII se cumplen, o sea, resulta un espacio lineal real.

Un ejemplo de espacio de infinitas dimensiones es también el conjunto de todas las funciones reales posibles de la variable real, entendiendo por suma de funciones y su producto por un número real lo que está convenido en la teoría de las funciones, es decir, como la suma y el producto por un número de los valores de las funciones para cada valor de la variable independiente.

Isomorfismo. Nuestro objetivo próximo consiste en la elección, entre todos los espacios lineales, de aquellos que naturalmente se pueden llamar espacios de dimensiones finitas. Introduzcamos primero un concepto general.

En la definición de espacio lineal se hablaba de las propiedades de las operaciones sobre los vectores, pero no se decía nada de las propiedades de los mismos vectores. En virtud de esto, puede ocurrir que, aunque los vectores de dos espacios lineales dados sean completamente distintos por su naturaleza, estos dos espacios no se distingan en nada desde el punto de vista de las propiedades de las operaciones. La definición exacta es:

Dos espacios lineales reales V y V' se llaman *isomorfos*, si entre los vectores de los mismos se ha establecido una correspondencia biunívoca —de modo que a cada vector a de V se asocia un vector a' de V', llamado **imagen** del vector a, teniendo que tener diferentes vectores de V diferentes imágenes y teniendo que ser cada vector de V' la imagen de cierto vector de V— y si en esta correspondencia, **la imagen de la suma de dos vectores es la suma de las imágenes de los mismos**

$$(a+b)'=a'+b', \tag{2}$$

y la imagen del producto de un vector por un número es el producto de la imagen de este vector por este mismo número,

$$(\alpha a)'=\alpha a'. \tag{3}$$

Señalemos, que la correspondencia biunívoca entre los espacios V y V' que satisface a las condiciones (2) y (3), se llama *correspondencia de isomorfismo.*

Así, pues, el espacio de los vectores-segmentos en el plano, que parten del origen de coordenadas, es isomorfo al espacio vectorial de dos dimensiones formado por pares ordenados de números reales: se obtiene una correspondencia de isomorfismo entre estos espacios, si en el plano se fija un sistema de coordenadas y a cada vector-segmento se asocia el par ordenado de sus coordenadas.

Demostremos la siguiente propiedad del isomorfismo de los espacios lineales:

en una correspondencia de isomorfismo entre los espacios V y V', la imagen del cero del espacio V es el cero del espacio V'.

En efecto, sea a un vector de V y sea a' su imagen en V'. Entonces, en virtud de (2),

$$a'=(a+0)'=a'+0',$$

es decir, $0'$ es el cero del espacio V'.

§ 30. Espacios de dimensiones finitas. Bases

Como fácilmente puede comprobar el lector, las dos definiciones de **dependencia lineal** de los vectores-filas que se hicieron en el § 9, así como la demostración de su equivalencia, empleaban

solamente las operaciones con los vectores y, por lo tanto, se pueden generalizar para el caso de espacios lineales cualesquiera. Por esta razón, en los espacios lineales definidos axiomáticamente se puede hablar de sistemas de vectores linealmente independientes, de sistemas linealmente independientes máximales (en caso de que existiesen), etc.

Si los espacios lineales V y V' son isomorfos, entonces un sistema de vectores $a_1, a_2, \ldots, a_k$ de V es linealmente dependiente si, y solo si, es linealmente dependiente el sistema de sus imágenes $a'_1, a'_2, \ldots, a'_k$ en V'.

Obsérvese, que si la correspondencia $a \rightarrow a'$ (para todos los a de V) es una correspondencia de isomorfismo entre V y V', entonces la correspondencia inversa $a' \rightarrow a$ también es de isomorfismo. Por esto, es suficiente examinar el caso en que el sistema $a_1, a_2, \ldots, a_k$ sea linealmente dependiente. Supongamos que existen unos números $\alpha_1, \alpha_2, \ldots, \alpha_k$, no simultáneamente iguales a cero, tales que

$$\alpha_1 a_1 + \alpha_2 a_2 + \ldots + \alpha_k a_k = 0.$$

Como ya sabemos, la imagen del segundo miembro de esta igualdad en el isomorfismo considerado es el cero $0'$ del espacio V'. Tomando la imagen del primer miembro y aplicando unas cuantas veces (2) y (3), se obtiene:

$$\alpha_1 a'_1 + \alpha_2 a'_2 + \ldots + \alpha_k a'_k = 0',$$

o sea, el sistema $a'_1, a'_2, \ldots, a'_k$ resulta también linealmente dependiente.

Espacios de dimensiones finitas. Se dice que un espacio lineal V es de *dimensión finita*, si se puede hallar en él un sistema finito de vectores linealmente independiente maximal; cualquier sistema tal de vectores se denominará *base* del espacio V.

Un espacio lineal de dimensión finita puede poseer muchas bases diversas. Así, en el espacio de vectores-segmentos en el plano, cualquier par de vectores, diferentes de cero y no situados en una recta, forma una base. Obsérvese, que nuestra definición de espacio de dimensión finita no responde a la pregunta si pueden existir o no en este espacio bases, compuestas de diferente número de vectores. Incluso, se podría suponer que en algunos espacios de dimensión finita existan bases con un número arbitrariamente grande de vectores. Ahora aclararemos la situación real existente.

Supongamos que el espacio lineal V posee una base

$$e_1, e_2, \ldots, e_n, \tag{1}$$

compuesta de n vectores. Si a es un vector arbitrario de V, como (1) es un sistema linealmente independiente maximal, a se expresa

linealmente mediante este sistema:

$$a = \alpha_1 e_1 + \alpha_2 e_2 + \ldots + \alpha_n e_n. \qquad (2)$$

Por otra parte, como el sistema (1) es linealmente independiente, la expresión (2) del vector a es única:
si

$$a = \alpha'_1 e_1 + \alpha'_2 e_2 + \ldots + \alpha'_n e_n,$$

se tiene

$$(\alpha_1 - \alpha'_1) e_1 + (\alpha_2 - \alpha'_2) e_2 + \ldots + (\alpha_n - \alpha'_n) e_n = 0,$$

de donde

$$\alpha_i = \alpha'_i, \qquad i = 1, 2, \ldots, n.$$

Por lo tanto, al vector a le corresponde unívocamente la fila

$$(\alpha_1, \alpha_2, \ldots, \alpha_n) \qquad (3)$$

de los coeficientes de su expresión (2) mediante la base (1) o, como vamos a decir, la *fila de sus coordenadas en la base* (1). Recíprocamente, toda fila de la forma (3), o sea, todo vector de n dimensiones en el sentido del cap. 2, es una fila de coordenadas en la base (1) para cierto vector del espacio V, precisamente para el vector que se expresa en la forma (2) mediante la base (1).

Por consiguiente, hemos obtenido una correspondencia biunívoca entre todos los vectores del espacio V y todos los vectores del espacio vectorial de filas de n dimensiones. Demostremos que esta correspondencia que, naturalmente, depende de la elección de la base (1), es una correspondencia de isomorfismo.

Tomemos también en el espacio V, además del vector a, que se expresa mediante la base (1) en la forma (2), un vector b, cuya expresión mediante la base (1) sea

$$b = \beta_1 e_1 + \beta_2 e_2 + \ldots + \beta_n e_n.$$

Entonces

$$a + b = (\alpha_1 + \beta_1) e_1 + (\alpha_2 + \beta_2) e_2 + \ldots + (\alpha_n + \beta_n) e_n,$$

es decir, *a la suma de los vectores a y b le corresponde la suma de las filas de sus coordenadas en la base* (1). Por otra parte,

$$\gamma a = (\gamma \alpha_1) e_1 + (\gamma \alpha_2) e_2 + \ldots + (\gamma \alpha_n) e_n,$$

o sea, *al producto de un vector a por un número γ le corresponde el producto de la fila de sus coordenadas en la base* (1) *por este mismo número γ.*

Con esto, queda demostrado el teorema siguiente:

Todo espacio lineal que posee una base de n vectores es isomorfo a un espacio vectorial de filas de n dimensiones.

Como ya sabemos, en una correspondencia de isomorfismo entre los espacios lineales, a un sistema de vectores linealmente depen-

diente le corresponde otro sistema linealmente dependiente, y viceversa; por lo tanto, a un sistema linealmente independiente le corresponde otro sistema linealmente independiente. De aquí se deduce que *en una correspondencia de isomorfismo, a la base le corresponde una base.*

En efecto, supongamos que en una correspondencia de isomorfismo entre los espacios V y V', a la base $e_1, e_2, \ldots, e_n$ del espacio V le corresponde el sistema de vectores $e'_1, e'_2, \ldots, e'_n$ del espacio V' que, aunque sea linealmente independiente, no es maximal. Por consiguiente, en V' se puede hallar un vector f' tal, que el sistema de vectores $e'_1, e'_2, \ldots, e'_n, f'$ se mantenga linealmente independiente. Sin embargo, en el isomorfismo considerado, el vector f' es la imagen de cierto vector f de V. Resulta, entonces, que el sistema de vectores $e_1, e_2, \ldots, e_n, f$ tiene que ser linealmente independiente, lo que contradice a la definición de la base.

Ya sabemos, que en el espacio vectorial de filas de n dimensiones (véase § 9), todos los sistemas linealmente independientes maximales constan de n vectores, que cualquier sistema de $n+1$ vectores es linealmente dependiente y que cualquier sistema de vectores linealmente independiente está contenido en un sistema linealmente independiente maximal. Aplicando las propiedades de las correspondencias de isomorfismo, establecidas anteriormente, llegamos a los resultados siguientes.

Todas las bases de un espacio lineal V de dimensión finita constan de un mismo número de vectores. Si este número es igual a n, V se llama *espacio lineal de n dimensiones* y el número n, *dimensión* de este espacio.

Todo sistema de $n+1$ vectores de un espacio lineal de n dimensiones es linealmente dependiente.

Todo sistema de vectores linealmente independiente de un espacio lineal de n dimensiones está contenido en una base de este espacio.

Ahora, es fácil comprobar que los ejemplos de espacios lineales reales indicados anteriormente, el espacio de sucesiones y el espacio de funciones, no son espacios de dimensión finita, pues, en cada uno de ellos el lector hallará sin dificultad sistemas linealmente independientes que constan de un número de vectores arbitrariamente grande.

Relación entre las bases. El objetivo de nuestro estudio son los espacios lineales de dimensión finita. Se entiende que al estudiar los espacios lineales de n dimensiones, se estudia realmente, el espacio vectorial de filas de n dimensiones que se introdujo en el cap. 2. Sin embargo, en este espacio se había elegido antes una base, en la que todos los vectores del espacio se determinaban por las filas de sus coordenadas; precisamente la base compuesta por los vectores unitarios, o sea, por los vectores que tienen una coordenada

igual a la unidad y todas las demás iguales a cero; ahora, todas las bases del espacio son para nosotros equivalentes.

Veamos la cantidad de bases que se pueden hallar en el espacio lineal de n dimensiones y cómo están ligadas estas bases entre sí.

Supongamos que en el espacio lineal V de n dimensiones se han dado las bases

$$e_1, e_2, \ldots, e_n \tag{4}$$

y

$$e'_1, e'_2, \ldots, e'_n. \tag{5}$$

Cada vector de la base (5), del mismo modo que cada vector del espacío V, se expresa unívocamente mediante la base (4),

$$e'_i = \sum_{j=1}^{n} \tau_{ij} e_j, \qquad i = 1, 2, \ldots, n. \tag{6}$$

La matriz

$$T = \begin{pmatrix} \tau_{11} & \ldots & \tau_{1n} \\ \cdot & \cdot\cdot\cdot & \cdot \\ \tau_{n1} & \ldots & \tau_{nn} \end{pmatrix},$$

cuyas filas son filas de las coordenadas de los vectores (5) en la base (4), se denomina *matriz de cambio* de la base (4) por la base (5).

En virtud de (6), la relación entre las bases (4) y (5) y la matriz de cambio T se puede expresar en forma de una igualdad matricial:

$$\begin{pmatrix} e'_1 \\ e'_2 \\ \cdot \\ \cdot \\ \cdot \\ e'_n \end{pmatrix} = \begin{pmatrix} \tau_{11} & \tau_{12} & \ldots & \tau_{1n} \\ \tau_{21} & \tau_{22} & \ldots & \tau_{2n} \\ \cdot & \cdot & \cdot\cdot\cdot & \cdot \\ \tau_{n1} & \tau_{n2} & \ldots & \tau_{nn} \end{pmatrix} \begin{pmatrix} e_1 \\ e_2 \\ \cdot \\ \cdot \\ \cdot \\ e_n \end{pmatrix} \tag{7}$$

o, en la forma:

$$e' = Te.$$

donde con e y e', se han designado, respectivamente, las bases (4) y (5) escritas en columna.

Por otra parte, si T' es la matriz de cambio de la base (5) por la base (4), se tiene

$$e = T'e'.$$

De aquí

$$e = (T'T)\, e,$$

$$e' = (TT')\, e',$$

y, como las bases e y e' son linealmente independientes, resulta

$$T'T = TT' = E,$$

de donde

$$T' = T^{-1}.$$

Con esto, queda demostrado que *la matriz de cambio de una base por otra es siempre una matriz no degenerada.*

Toda matriz cuadrada no degenerada de orden n con elementos reales es la matriz de cambio de una base dada del espacio lineal real de n dimensiones por otra base.

En efecto, supongamos dada la base (4) y la matriz T, de orden n, no degenerada. Tomemos por (5) el sistema de vectores, para los que las filas de la matriz T son filas de coordenadas en la base (4); por consiguiente, se cumple la igualdad (7). Los vectores (5) son linealmente independientes, puesto que la dependencia lineal entre ellos daría lugar a la dependencia lineal de las filas de la matriz T, lo cual es absurdo, pues T no es degenerada. En consecuencia, el sistema (5), siendo linealmente independiente y constando de n vectores, es una base de nuestro espacio y la matriz T es la matriz de cambio de la base (4) por la base (5).

Vemos, pues, que en el espacio lineal de n dimensiones se pueden hallar tantas bases diversas, cuantas matrices cuadradas diversas no degeneradas de orden n existan. Claro que, en este caso, dos bases que consten de los mismos vectores, pero escritos en orden diverso, se consideran diferentes.

Transformación de las coordenadas de un vector. Supongamos que en el espacio lineal de n dimensiones se han dado las bases (4) y (5) con la matriz de cambio $T=(\tau_{ij})$,

$$e'=Te.$$

Hallemos la relación existente entre las filas de coordenadas de un vector arbitrario a en estas bases.

Supongamos que

$$a=\sum_{j=1}^{n}\alpha_j e_j, \tag{8}$$

$$a=\sum_{i=1}^{n}\alpha'_i e'_i.$$

Aplicando (6), resulta:

$$a=\sum_{i=1}^{n}\alpha'_i\left(\sum_{j=1}^{n}\tau_{ij}e_j\right)=\sum_{j=1}^{n}\left(\sum_{i=1}^{n}\alpha'_i\tau_{ij}\right)e_j.$$

Comparando con (8) y aplicando la unicidad de la expresión de un vector mediante la base, se obtiene:

$$\alpha_j=\sum_{i=1}^{n}\alpha'_i\tau_{ij},\qquad j=1,\ 2,\ \ldots,\ n,$$

o sea, se cumple la igualdad matricial:

$$(\alpha_1,\ \alpha_2,\ \ldots,\ \alpha_n)=(\alpha'_1,\ \alpha'_2,\ \ldots,\ \alpha'_n)\,T.$$

Por lo tanto, *la fila de coordenadas de un vector a en la base e es igual a la fila de coordenadas de este vector en la base e', multiplicada a la derecha por la matriz de cambio de la base e por la base e'.*

Naturalmente, de aquí se deduce la igualdad

$$(\alpha_1', \alpha_2', \ldots, \alpha_n') = (\alpha_1, \alpha_2, \ldots, \alpha_n) T^{-1}.$$

Ejemplo. Examinemos el espacio lineal real de tres dimensiones con la base

$$e_1, e_2, e_3. \tag{9}$$

Los vectores

$$\left.\begin{aligned} e_1' &= 5e_1 - e_2 - 2e_3, \\ e_2' &= 2e_1 + 3e_2, \\ e_3' &= -2e_1 + e_2 + e_3 \end{aligned}\right\} \tag{10}$$

también forman una base en este espacio, siendo

$$T = \begin{pmatrix} 5 & -1 & -2 \\ 2 & 3 & 0 \\ -2 & 1 & 1 \end{pmatrix},$$

la matriz de cambio de (9) por (10); de donde

$$T^{-1} = \begin{pmatrix} 3 & -1 & 6 \\ -2 & 1 & -4 \\ 8 & -3 & 17 \end{pmatrix}.$$

Por esto, el vector

$$a = e_1 + 4e_2 - e_3$$

tiene en la base (10) la fila de coordenadas

$$(\alpha_1', \alpha_2', \alpha_3') = (1, 4, -1) \begin{pmatrix} 3 & -1 & 6 \\ -2 & 1 & -4 \\ 8 & -3 & 17 \end{pmatrix} = (-13, 6, -27),$$

o sea,

$$a = -13e_1' + 6e_2' - 27e_3'.$$

§ 31. Transformaciones lineales

Ya nos encontramos en el cap. 3 con el concepto de transformación lineal de las indeterminadas. El concepto que se va a introducir ahora lleva el mismo nombre, pero tiene diferente carácter. Ahora bien, se pueden indicar ciertas relaciones entre estos dos conceptos homónimos.

Sea dado un espacio lineal real de n dimensiones, que lo designaremos con V_n. Examinemos una *transformación* de este espacio, o sea, una correspondencia que asocia **a cada** vector a del espacio V_n cierto vector a' de este mismo espacio. El vector a' se llama *imagen* del vector a en la transformación considerada.

Si la transformación se designa con φ, convendremos en designar la imagen del vector a con $a\varphi$ y no con $\varphi(a)$ o φa, como es usual para

el lector. Por lo tanto

$$a' = a\varphi.$$

Una transformación φ del espacio lineal V_n se denomina *transformación lineal* de este espacio, si ésta transforma la suma de dos vectores cualesquiera a, b en la suma de las imágenes de estos vectores,

$$(a + b)\varphi = a\varphi + b\varphi, \tag{1}$$

y el producto de cualquier vector a por cualquier número α, en el producto de la imagen del vector a por este mismo número α,

$$(\alpha a)\varphi = \alpha(a\varphi). \tag{2}$$

De esta definición resulta inmediatamente que *la transformación lineal del espacio lineal transforma cualquier combinación lineal de los vectores dados $a_1, a_2, \ldots, a_k$ en la combinación lineal (con los mismos coeficientes) de las imágenes de estos vectores*:

$$(\alpha_1 a_1 + \alpha_2 a_2 + \ldots + \alpha_k a_k)\varphi = \alpha_1(a_1\varphi) + \alpha_2(a_2\varphi) + \ldots + \alpha_k(a_k\varphi). \tag{3}$$

Demostremos la siguiente afirmación:

Para cualquier transformación lineal φ del espacio lineal V_n, el vector nulo 0 se mantiene inmóvil,

$$0\varphi = 0,$$

y la imagen del vector opuesto al vector dado a, es el vector opuesto a la imagen del vector a,

$$(-a)\varphi = -a\varphi.$$

En efecto, si b es un vector arbitrario, en virtud de (2), se tiene

$$0\varphi = (0 \cdot b)\varphi = 0 \cdot (b\varphi) = 0.$$

Por otra parte,

$$(-a)\varphi = [(-1)a]\varphi = (-1)(a\varphi) = -a\varphi.$$

El concepto de transformación lineal de un espacio lineal surgió como una generalización de la transformación afín del plano o del espacio de tres dimensiones, tratada en el curso de geometría analítica; en efecto, las condiciones (1) y (2) para las transformaciones afines se cumplen. Estas condiciones también se cumplen para las proyecciones de los vectores en el plano o para las proyecciones sobre una recta (o sobre un plano) en el espacio de tres dimensiones. Por lo tanto, en el espacio lineal de dos dimensiones, de vectores-segmentos que parten del origen de coordenadas en el plano, la transformación de cualquier vector en su proyección sobre un eje que pase por el origen de coordenadas, es una transformación lineal.

Son ejemplos de transformaciones lineales en un espacio arbitrario V_n, la *transformación idéntica* ε, que mantiene a cada vector a en su sitio,

$$a\varepsilon = a,$$

y la *transformación nula* ω, que transforma cualquier vector a en el vector nulo,

$$a\omega = 0.$$

Estudiemos ahora todas las transformaciones lineales del espacio V_n. Sea

$$e_1,\ e_2,\ \ldots,\ e_n \tag{4}$$

una base de este espacio; igual que antes, la base (4), colocada en columna, se designará con e. Como cualquier vector a del espacio V_n se representa unívocamente en forma de una combinación lineal de los vectores de la base (4), la imagen del vector a, en virtud de (3), se expresa mediante las imágenes de los vectores (4) con los mismos coeficientes. En otras palabras, *toda transformación lineal* φ *del espacio* V_n *se determina unívocamente por las imágenes* $e_1\varphi$, $e_2\varphi$, ..., $e_n\varphi$ *de todos los vectores de una base* (4) *fijada.*

Cualquiera que sea el sistema de n *vectores ordenados*

$$c_1,\ c_2,\ \ldots,\ c_n, \tag{5}$$

del espacio V_n, *existe una transformación lineal de este espacio y sólo una, tal que el sistema* (5) *es el sistema de imágenes de los vectores de la base* (4) *en esta transformación,*

$$e_1\varphi = c_i, \qquad i = 1,\ 2,\ \ldots,\ n. \tag{6}$$

La unicidad de la transformación φ se demostró anteriormente y sólo queda por demostrar su existencia. Determinemos una transformación φ del modo siguiente: si a es un vector arbitrario del espacio y

$$a = \sum_{i=1}^{n} \alpha_i e_i$$

es su expresión en la base (4), supondremos que

$$a\varphi = \sum_{i=1}^{n} \alpha_i c_i. \tag{7}$$

Demostremos que esta transformación es lineal. Si

$$b = \sum_{i=1}^{n} \beta_i e_i$$

es cualquier otro vector del espacio, se tiene

$$(a+b)\varphi=\left[\sum_{i=1}^{n}(\alpha_i+\beta_i)e_i\right]\varphi=\sum_{i=1}^{n}(\alpha_i+\beta_i)c_i=$$
$$=\sum_{i=1}^{n}\alpha_i c_i+\sum_{i=1}^{n}\beta_i c_i=a\varphi+b\varphi.$$

Si γ es un número cualquiera, entonces

$$(\gamma a)\varphi=\left[\sum_{i=1}^{n}(\gamma a_i)e_i\right]\varphi=\sum_{i=1}^{n}(\gamma\alpha_i)c_i=\gamma\sum_{i=1}^{n}\alpha_i c_i=\gamma(a\varphi).$$

En lo que se refiere a la igualdad (6), ésta se cumple debido a la definición (7) de la transformación φ, puesto que todas la coordenadas del vector e_i en la base (4) son iguales a cero, excepto la i-ésima coordenada, que es igual a la unidad.

Por consiguiente, hemos establecido una correspondencia biunívoca entre todas las transformaciones del espacio lineal V_n y todos los sistemas ordenados (5) *formados por n vectores de este espacio.*

Sin embargo, todo vector c_i posee una determinada expresión en la base (4),

$$c_i=\sum_{j=1}^{n}\alpha_{ij}e_j,\qquad i=1,\ 2,\ \ldots,\ n. \tag{8}$$

Con las coordenadas del vector c_i en la base (4) se puede formar una matriz cuadrada

$$A=(\alpha_{ij}) \tag{9}$$

tomando por i-ésima fila la fila de coordenadas del vector c_i, $i=1,\ 2,\ \ldots,\ n$. Como el sistema (5) es arbitrario, la matriz A será una matriz cuadrada arbitraria de orden n con elementos reales.

Por lo tanto, resulta una correspondencia biunívoca entre todas las transformaciones lineales del espacio V_n y todas las matrices cuadradas de orden n; por supuesto, esta correspondencia depende de la elección de la base (4).

Se dice que la matriz A *determina* la transformación lineal φ en la base (4), o, abreviadamente, que A es *la matriz de la transformación lineal* φ en la base (4). Si designamos con $e\varphi$ la columna formada por las imágenes de los vectores de la base (4), entonces, de (6), (8) y (9) se deduce la siguiente igualdad matricial, que describe totalmente la relación existente entre la transformación lineal φ, la base e y la matriz A que determina esta transformación lineal en esta base:

$$e\varphi=Ae. \tag{10}$$

He aquí cómo se hallan las coordenadas de la imagen $a\varphi$ del vector a en la base (4), conociendo las coordenadas del mismo vector

a en la misma base y la matriz A de la transformación lineal φ. Si

$$a=\sum_{i=1}^{n}\alpha_i e_i,$$

se tiene

$$a\varphi=\sum_{i=1}^{n}\alpha_i (e_i\varphi),$$

lo que es equivalente a la igualdad matricial

$$a\varphi=(\alpha_1,\ \alpha_2,\ \ldots,\ \alpha_n)(e\varphi).$$

Aplicando (10) y teniendo en cuenta que la multiplicación de las matrices es asociativa también cuando una de las matrices es una columna formada por vectores (lo cual fácilmente se comprueba), resulta:

$$a\varphi=[(\alpha_1,\ \alpha_2,\ \ldots,\ \alpha_n)A]\,e.$$

De aquí se deduce que *la fila de coordenadas del vector* $a\varphi$ *es igual a la fila de coordenadas del vector* a, *multiplicada a la derecha por la matriz* A *de la transformación lineal, todo efectuado en la base* (4).

Ejemplo. Supongamos que en la base e_1, e_2, e_3 del espacio lineal de tres dimensiones, la transformación lineal se da mediante la matriz

$$A=\begin{pmatrix}-2 & 1 & 0\\ 1 & 3 & 2\\ 0 & -4 & 1\end{pmatrix}.$$

Si

$$a=5e_1+e_2-2e_3,$$

entonces

$$(5,\ 1,\ -2)\begin{pmatrix}-2 & 1 & 0\\ 1 & 3 & 2\\ 0 & -4 & 1\end{pmatrix}=(-9,\ 16,\ 0),$$

o sea,

$$a\varphi=-9e_1+16e_2.$$

Relación entre las matrices de una transformación lineal en diversas bases. Naturalmente, la matriz que determina la transformación lineal depende de la elección de la base. Hallemos la relación entre las matrices que determinan una misma transformación lineal pero en bases diferentes.

Sean dadas las bases e y e' con la matriz de cambio T,

$$e'=Te, \tag{11}$$

y supongamos que la transformación lineal φ se determina en estas bases por las matrices A y A', respectivamente,

$$e\varphi=Ae,\quad e'\varphi=A'e'. \tag{12}$$

En virtud de (11), la segunda de las igualdades (12) da lugar a la igualdad

$$(Te)\varphi = A'(Te).$$

Pero

$$(Te)\varphi = T(e\varphi).$$

En efecto, si $(\tau_{i1}, \tau_{i2}, \ldots, \tau_{in})$ es la i-ésima fila de la matriz T, se tiene

$$(\tau_{i1}e_1 + \tau_{i2}e_2 + \ldots + \tau_{in}e_n)\varphi = \tau_{i1}(e_1\varphi) + \tau_{i2}(e_2\varphi) + \ldots + \tau_{in}(e_n\varphi).$$

Por lo tanto, en virtud de (12),

$$(Te)\varphi = T(e\varphi) = T(Ae) = (TA)e,$$
$$A'(Te) = (A'T)e,$$

o sea,

$$(TA)e = (A'T)e.$$

Si al menos para un i, $1 \leqslant i \leqslant n$, la i-ésima fila de la matriz TA fuese diferente de la i-ésima fila de la matriz $A'T$, entonces dos distintas combinaciones lineales de los vectores $e_1, e_2, \ldots, e_n$ resultarían iguales entre sí, lo cual es absurdo, puesto que la base e es linealmente independiente. Por lo tanto,

$$TA = A'T,$$

y como la matriz de cambio T no es degenerada, de aquí resulta que

$$A' = TAT^{-1}, \quad A = T^{-1}A'T. \tag{13}$$

Dos matrices, B y C, se llaman *semejantes*, si están ligadas por la igualdad

$$C = Q^{-1}BQ,$$

donde Q es una matriz no degenerada. En este caso, se dice que la matriz C es la *transformada* de la matriz B por la matriz Q.

Por lo tanto, las igualdades (13) demostradas anteriormente se pueden enunciar en forma del siguiente importante **teorema:**

Las matrices que determinan una misma transformación lineal en diferentes bases, son semejantes entre sí. Además, la matriz de la transformación lineal φ en la base e' se obtiene transformando la matriz de esta transformación en la base e por la matriz de cambio de la base e' a la base e.

Subrayemos que, si la matriz A determina la transformación φ en la base e, cualquier matriz B semejante a la matriz A,

$$B = Q^{-1}AQ,$$

también determina la transformación φ en cierta base, precisamente, en la base que se obtiene de la base e mediante la matriz de cambio Q^{-1}.

Operaciones con las transformaciones lineales. Como ya se demostró, asociando a cada transformación lineal del espacio V_n su matriz en una base fija, resulta una correspondencia biunívoca entre todas las transformaciones lineales y todas las matrices cuadradas de orden n. Es natural esperar que a las operaciones de adición y multiplicación de las matrices, y también a la multiplicación de una matriz por un número, les correspondan unas operaciones análogas con las transformaciones lineales.

Sean dadas en el espacio V_n las transformaciones φ y ψ. Llamemos *suma* de estas transformaciones a la transformación $\varphi+\psi$, determinada por la igualdad

$$a(\varphi+\psi)=a\varphi+a\psi; \tag{14}$$

por consiguiente, ésta transforma cualquier vector a en la suma de sus imágenes en las transformaciones φ y ψ.

La transformación $\varphi+\psi$ es lineal. En efecto, para cualesquiera vectores a y b y cualquier número α,

$$(a+b)(\varphi+\psi)=(a+b)\varphi+(a+b)\psi=$$
$$=a\varphi+b\varphi+a\psi+b\psi=a(\varphi+\psi)+b(\varphi+\psi);$$
$$(\alpha a)(\varphi+\psi)=(\alpha a)\varphi+(\alpha a)\psi=\alpha(a\varphi)+\alpha(a\psi)=$$
$$=\alpha(a\varphi+a\psi)=\alpha[a(\varphi+\psi)].$$

Por otra parte, llamemos *producto* de las transformaciones lineales φ y ψ a la transformación $\varphi\psi$ determinada por la igualdad

$$a(\varphi\psi)=(a\varphi)\psi, \tag{15}$$

es decir, que se obtiene como resultado de la realización sucesiva de las transformaciones φ y ψ.

La transformación $\varphi\psi$ es lineal:

$$(a+b)(\varphi\psi)=[(a+b)\varphi]\psi=(a\varphi+b\varphi)\psi=$$
$$=(a\varphi)\psi+(b\varphi)\psi=a(\varphi\psi)+b(\varphi\psi);$$
$$(\alpha a)(\varphi\psi)=[(\alpha a)\varphi]\psi=[\alpha(a\varphi)]\psi=\alpha[(a\varphi)\psi]=\alpha[a(\varphi\psi)].$$

Finalmente, llamemos *producto de la transformación lineal φ por el número $\varkappa$* a la transformación $\varkappa\varphi$ determinada por la igualdad

$$a(\varkappa\varphi)=\varkappa(a\varphi); \tag{16}$$

de aquí que en la transformación φ las imágenes de todos los vectores se multiplican por el número $\varkappa$.

La transformación $\varkappa\varphi$ es lineal:

$$(a+b)(\varkappa\varphi)=\varkappa[(a+b)\varphi]=\varkappa(a\varphi+b\varphi)=$$
$$=\varkappa(a\varphi)+\varkappa(b\varphi)=a(\varkappa\varphi)+b(x\varphi);$$
$$(\alpha a)(\varkappa\varphi)=\varkappa[(\alpha a)\varphi]=\varkappa[\alpha(a\varphi)]=\alpha[\varkappa(a\varphi)]=\alpha[a(\varkappa\varphi)].$$

Supongamos que en la base $e_1, e_2, \ldots, e_n$, las transformaciones φ y ψ se determinan por las matrices $A = (\alpha_{ij})$ y $B = (\beta_{ij})$, respectivamente,

$$e\varphi = Ae, \quad e\psi = Be.$$

Entonces, en virtud de (14),

$$e_i(\varphi + \psi) = e_i\varphi + e_i\psi = \sum_{j=1}^{n} \alpha_{ij}e_j + \sum_{j=1}^{n} \beta_{ij}e_j = \sum_{j=1}^{n} (\alpha_{ij} + \beta_{ij})\, e_j,$$

o sea

$$e(\varphi + \psi) = (A + B)\, e.$$

Por lo tanto, *la matriz de la suma de transformaciones lineales en cualquier base es igual a la suma de las matrices de estas transformaciones en esta misma base.*

Por otra parte, en virtud de (15),

$$e_i(\varphi\psi) = (e_i\varphi)\,\psi = \left(\sum_{j=1}^{n} \alpha_{ij}e_j\right)\psi = \sum_{j=1}^{n} \alpha_{ij}(e_j\psi) =$$

$$= \sum_{j=1}^{n} \alpha_{ij}\left(\sum_{k=1}^{n} \beta_{jk}e_k\right) = \sum_{k=1}^{n}\left(\sum_{j=1}^{n} \alpha_{ij}\beta_{jk}\right) e_k,$$

o sea,

$$e(\varphi\psi) = (AB)\, e.$$

En otras palabras, *la matriz del producto de transformaciones lineales en cualquer base es igual al producto de las matrices de estas transformaciones en la misma base.*

Finalmente, en virtud de (16),

$$e_i(\varkappa\varphi) = \varkappa(e_i\varphi) = \varkappa \sum_{j=1}^{n} \alpha_{ij}e_j = \sum_{j=1}^{n} (\varkappa\alpha_{ij})\, e_j,$$

o sea,

$$e(\varkappa\varphi) = (\varkappa A)\, e.$$

Por consiguiente, *la matriz que determina en cierta base el producto de la transformación lineal φ por el número $\varkappa$, es igual al producto de la matriz de la misma transformación φ en esta base por el número $\varkappa$.*

De los resultados obtenidos se deduce que las operaciones con las transformaciones lineales poseen las mismas propiedades que las operaciones con las matrices. Así, pues, la suma de transformaciones lineales es conmutativa y asociativa, y el producto es asociativo, aunque para $n > 1$ no es conmutativo. Para las transformaciones lineales existe la resta unívoca. Obsérvese también que *entre las trans-*

formaciones lineales, la transformación idéntica ε *desempeña el papel de la unidad, y la transformación nula* ω, *el papel del cero.* En efecto, en cualquier base, la transformación ε se determina por la matriz unidad, y la transformación ω, por la matriz nula.

§ 32. Subespacios lineales

Un subconjunto L del espacio lineal V se llama *subespacio lineal* de este espacio, si él mismo es un espacio lineal con respecto a las operaciones de suma de vectores y de multiplicación de un vector por un número, determinadas en V. Así, pues, en el espacio euclídeo de tres dimensiones, el conjunto de vectores que parten del origen de coordenadas y que están situados en un plano (o en una recta) que pasa por el origen, es un subespacio lineal.

Para que un subconjunto no vacío L del espacio V sea un subespacio lineal de éste, es suficiente que se cumplan las condiciones siguientes:

1. *Si los vectores a y b pertenecen a L, el vector* $a + b$ *también pertenece a L.*

2. *Si el vector a pertenece a L, el vector* αa *también pertenece a L para cualquier valor del número* α.

En efecto, en virtud de la condición 2, el conjunto L contiene el vector nulo, pues, si el vector a pertenece a L, el vector $0 \cdot a = 0$ también pertenece a L. Luego, junto con cada uno de sus vectores a, y otra vez en virtud de la condición 2, el vector opuesto $-a = (-1) \cdot a$ también pertenece a L, por lo cual, debido a la condición 1, también pertenece a L la diferencia de dos vectores cualesquiera de L. En lo que se refiere a las demás condiciones incluidas en la definición del espacio lineal, cumpliéndose éstas en V, también se cumplen en L.

Pueden servir de ejemplos de subespacios lineales del espacio V, el mismo espacio V, así como el conjunto compuesto del solo vector nulo, denominado *subespacio nulo*. De mayor interés es el siguiente: tomemos en el espacio V cualquier sistema finito de vectores

$$a_1, a_2, \ldots, a_r \qquad (1)$$

y designemos con L el conjunto de todos los vectores que son combinaciónes lineales de los vectores (1). Demostremos que L es un subespacio lineal. En efecto, si

$$b = \alpha_1 a_1 + \alpha_2 a_2 + \ldots + \alpha_r a_r, \quad c = \beta_1 a_1 + \beta_2 a_2 + \ldots + \beta_r a_r,$$

se tiene

$$b + c = (\alpha_1 + \beta_1) a_1 + (\alpha_2 + \beta_2) a_2 + \ldots + (\alpha_r + \beta_r) a_r,$$

o sea, el vector $b+c$ pertenece a L; también pertenece a L el vector

$$\gamma b = (\gamma\alpha_1)\, a_1 + (\gamma\alpha_2)\, a_2 + \ldots + (\gamma\alpha_r)\, a_r$$

para cualquier número γ.

Suele decirse que este subespacio lineal L *está engendrado* por el sistema de vectores (1); en particular, los mismos vectores (1) pertenecen a L.

Por cierto, *todo subespacio lineal de un espacio lineal de dimensión finita se engendra por un sistema finito de vectores*, ya que, si el subespacio no es el nulo, posee incluso una base finita. La dimensión del subespacio lineal L no es mayor que la dimensión n del mismo espacio V_n, siendo igual a n solamente cuando $L = V_n$. Evidentemente, la dimensión del subespacio nulo se debe tomar igual a cero.

Para cualquier k, $0 < k < n$, en el espacio V_n existen subespacios lineales de dimensión k; para esto, es suficiente considerar el subespacio engendrado por cualquier sistema de k vectores linealmente independientes.

Supongamos que en el espacio V se han dado los subespacios lineales L_1 y L_2. Como fácilmente se comprueba, el conjunto de vectores L_0 que pertenecen simultáneamente a L_1 y a L_2, es un subespacio lineal; éste se llama *intersección* de los subespacios L_1 y L_2. Por otra parte, también es un subespacio lineal la suma $\overline{L}$ de los subespacios L_1 y L_2, o sea, el conjunto de todos los vectores de V que se representan en forma de una suma de dos vectores, uno de los cuales pertenece a L_1 y el otro, a L_2. Si d_1, d_2, d_0 y $\overline{d}$ son las dimensiones respectivas de los subespacios L_1, L_2, L_0 y $\overline{L}$, se cumple la igualdad:

$$\overline{d} = d_1 + d_2 - d_0, \tag{2}$$

es decir, *la dimensión de la suma de dos subespacios es igual a la suma de las dimensiones de estos subespacios menos la dimensión de su intersección.*

Para la demostración, se toma una base arbitraria

$$a_1,\ a_2,\ \ldots,\ a_{d_0} \tag{3}$$

del subespacio L_0 y se completa hasta que se forme una base

$$a_1,\ a_2,\ \ldots,\ a_{d_0},\ b_{d_0+1},\ \ldots,\ b_{d_1} \tag{4}$$

del subespacio L_1, y hasta que se forme una base

$$a_1,\ a_2,\ \ldots,\ a_{d_0},\quad c_{d_0+1},\ \ldots,\ c_{d_2} \tag{5}$$

del subespacio L_2. Aplicando la definición del subespacio $\overline{L}$, se observa sin dificultad que éste se engendra por el sistema de vectores

$$a_1,\ a_2,\ \ldots,\ a_{d_0},\ b_{d_0+1},\ \ldots,\ b_{d_1},\ c_{d_0+1},\ \ldots,\ c_{d_2}. \tag{6}$$

Por consiguiente, la fórmula (2) quedará demostrada si se demuestra que el sistema (6) es linealmente independiente.

Supongamos que se cumple la igualdad

$$\alpha_1 a_1 + \alpha_2 a_2 + \ldots + \alpha_{d_0} a_{d_0} + \beta_{d_0+1} b_{d_0+1} + \ldots + \beta_{d_1} b_{d_1} + \\ + \gamma_{d_0+1} c_{d_0+1} + \ldots + \gamma_{d_2} c_{d_2} = 0$$

con ciertos coeficientes numéricos. Entonces,

$$d = \alpha_1 a_1 + \alpha_2 a_2 + \ldots + \alpha_{d_0} a_{d_0} + \beta_{d_0+1} b_{d_0+1} + \ldots + \beta_{d_1} b_{d_1} = \\ = -\gamma_{d_0+1} c_{d_0+1} - \ldots - \gamma_{d_2} c_{d_2}. \quad (7)$$

El primer miembro de esta igualdad pertenece a L_1, el segundo, a L_2. Por consiguiente, el vector d, que es igual tanto al primer miembro como al segundo miembro de esta igualdad, pertenece a L_0, por lo cual se expresa linealmente mediante la base (3). Sin embargo, el segundo miembro de la igualdad (7) muestra que el vector d también se expresa linealmente mediante los vectores $c_{d_0+1}, \ldots, c_{d_2}$. Como el sistema (5) es linealmente independiente, resulta que todos los coeficientes $\gamma_{d_0+1}, \ldots, \gamma_{d_2}$ son iguales a cero, es decir, $d = 0$, y por lo tanto, como el sistema (4) es linealmente independiente, también son iguales a cero todos los coeficientes $\alpha_1, \ldots, \alpha_{d_0}, \beta_{d_0+1}, \ldots, \beta_{d_1}$. Con esto, queda demostrada la independencia lineal del sistema (6).

Se recomienda al lector comprobar que nuestra demostración es válida también para el caso en que el subespacio L_0 sea nulo, o sea, $d_0 = 0$.

Campo de valores y núcleo de una transformación lineal. Supongamos que en el espacio lineal V_n es dada una transformación lineal φ. De las definiciones de subespacio lineal y de tranformación lineal se deduce inmediatamente que si L es un subespacio lineal cualquiera del espacio V_n, **el conjunto $L\varphi$ de las imágenes de todos los vectores de L en la transformación φ también es un subespacio lineal**; en particular, el conjunto $V_n\varphi$ de las imágenes de todos los vectores del espacio V_n es también un subespacio lineal; este conjunto se llama *campo de valores* de la transformación φ.

Hallemos la dimensión del campo de valores. Para esto, obsérvese que, como todas las matrices que determinan la transformación φ en diferentes bases son semejantes entre sí, todas ellas tienen un mismo rango en virtud del último teorema del § 14. Por consiguiente, este número puede recibir el nombre de *rango* de la transformación lineal φ.

La dimensión del campo de valores de una transformación lineal φ es igual al rango de esta última.

En efecto, supongamos que φ *se determina en la base* e_1, $e_2, \ldots, e_n$ por la matriz A. El subespacio $V_n\varphi$ se engendra por los vectores

$$e_1\varphi,\ e_2\varphi,\ \ldots,\ e_n\varphi \tag{8}$$

y, en particular, cualquier subsistema linealmente independiente, maximal del sistema (8) será base del subespacio $V_n\varphi$. Pero, el maximo número de vectores linealmente independientes del sistema (8) es igual al máximo número de filas linealmente independientes de la matriz A, es decir, es igual al rango de esta matriz. El teorema queda demostrado.

Ya sabemos que en una transformación lineal φ, el vector nulo *se transforma en sí mismo. El conjunto* $N(\varphi)$ *de todos los vectores* del espacio V_n que se transforman en el vector nulo en la transformación φ, no es, pues, vacío, y es, evidentemente, un subespacio lineal. Este subespacio se denomina *núcleo* de la transformación φ, y su dimensión lleva el nombre de *defecto* de la misma.

Para cualquier transformación lineal φ *del espacio* V_n, *la suma del rango y del defecto de la misma es igual a la dimensión* n *de todo el espacio.*

En efecto, si r es el rango de la transformación φ, el subespacio $V_n\varphi$ posee una base de r vectores

$$a_1,\ a_2,\ \ldots,\ a_r. \tag{9}$$

En el espacio V_n se pueden elegir tales vectores

$$b_2,\ b_2,\ \ldots,\ b_r, \tag{10}$$

que

$$b_i\varphi = a_i, \quad i = 1,\ 2,\ \ldots,\ r;$$

evidentemente, la elección de los vectores (10) no es única. Si alguna combinación lineal no trivial de los vectores (10) se transformase en cero y, en particular, si los vectores (10) fuesen linealmente dependientes, los vectores (9) también resultarían linealmente dependientes, en contra de la suposición. Por esto, el subespacio lineal L engendrado por los vectores (10), tiene la dimensión r, y su intersección con el subespacio $N(\varphi)$ es igual a cero.

Por otra parte, la suma de los subespacios L y $N(\varphi)$ coincide con todo el espacio V_n. En efecto, siendo c cualquier vector del espacio, el vector $d = c\varphi$ pertenece, naturalmente, al subespacio $V_n\varphi$. Entonces, existe en el subespacio L un vector b tal que

$$b\varphi = d;$$

el vector b se expresa mediante el sistema (10) con los mismos coeficientes con que se expresa el vector d mediante la base (9). Por consiguiente

$$c = b + (c - b).$$

donde el vector $c-b$ pertenece al subespacio $N(\varphi)$, puesto que

$$(c-b)\varphi = c\varphi - b\varphi = d - d = 0.$$

La afirmación del teorema se deduce de los resultados obtenidos y de la fórmula (2).

Transformaciones lineales no degeneradas. Una transformación lineal φ del espacio lineal V_n se llama *no degenerada*, si se satisface cualquiera de las condiciones siguientes, cuya equivalencia es consecuencia inmediata de los teoremas demostrados anteriormente:

1. **El rango de la transformación φ es igual a n.**
2. **El campo de valores de la transformación φ es todo el espacio V_n.**
3. **El defecto de la transformación φ es igual a cero.**

Para las transformaciones lineales no degeneradas se pueden dar también muchas otras definiciones equivalentes a las señaladas y, en particular, las definiciones 4—6 que siguen.

4. **Diferentes vectores del espacio V_n tienen en la transformación φ diferentes imágenes.**

En efecto, si la transformación φ posee la propiedad 4, el núcleo de esta transformación consta solamente del vector nulo, o sea, se cumple también la condición 3. Si los vectores a y b son tales que $a \neq b$, pero $a\varphi = b\varphi$, se tiene $a-b \neq 0$, pero $(a-b)\varphi = 0$, es decir, no se cumple la condición 3.

De 2 y 4 se deduce:

5. **La transformación φ es una correspondencia biunívoca del espacio V_n sobre todo este espacio.**

De 5 se deduce que, para una transformación lineal φ no degenerada, existe la *transformación inversa* φ^{-1} que transforma cada vector $a\varphi$ en el vector a,

$$(a\varphi)\varphi^{-1} = a.$$

La transformación φ^{-1} es lineal, puesto que

$$(a\varphi + b\varphi)\varphi^{-1} = [(a+b)\varphi]\varphi^{-1} = a+b,$$

$$[\alpha(a\varphi)]\varphi^{-1} = [(\alpha a)\varphi]\varphi^{-1} = \alpha a.$$

De la definición de la transformación φ^{-1}, se deduce que

$$\varphi\varphi^{-1} = \varphi^{-1}\varphi = \varepsilon; \tag{11}$$

las mismas igualdades (11) se pueden considerar como la definición de la transformación inversa. De aquí y de los últimos resultados del párrafo anterior, se deduce que *si una transformación lineal φ no degenerada se determina en cierta base por la matriz A, que no es degenerada en virtud de la propiedad* 1, *entonces la transformación φ^{-1} se determina en esta base por la matriz A^{-1}.*

Por lo tanto, llegamos a la siguiente definición de transformación lineal no degenerada:

6. **Para la transformación φ existe la transformación lineal inversa φ^{-1}.**

§ 33. Raíces características y valores propios

Sea $A = (\alpha_{ij})$ una matriz cuadrada de orden n con elementos reales. Sea, por otra parte, λ una indeterminada. Entonces la matriz $A-\lambda E$, donde E es la matriz unidad de orden n, se llama *matriz característica* de la matriz A. Como en la diagonal principal de la matriz λE, figura λ, siendo iguales a cero todos los demás elementos, resulta

$$A-\lambda E=\begin{pmatrix} \alpha_{11}-\lambda & \alpha_{12} & \dots & \alpha_{1n} \\ \alpha_{21} & \alpha_{22}-\lambda & \dots & \alpha_{2n} \\ \dots & \dots & \dots & \dots \\ \alpha_{n1} & \alpha_{n2} & \dots & \alpha_{nn}-\lambda \end{pmatrix}.$$

El determinante de la matriz $A-\lambda E$ es un polinomio en λ, de orden n. En efecto, el producto de los elementos que figuran en la diagonal principal es un polinomio en λ con el término superior $(-1)^n\lambda^n$. Todos los demás términos del determinante no contienen al menos dos de los elementos que figuran en la diagonal principal, por lo que su grado respecto a λ no es mayor que $n-2$. Los coeficientes de este polinomio se podrían hallar fácilmente. Así, pues, el coeficiente de λ^{n-1} es igual a $(-1)^{n-1}(\alpha_{11}+\alpha_{22}+\dots+\alpha_{nn})$ *y el término independiente coincide con el determinante de la matriz A.*

El polinomio de n-ésimo grado $|A-\lambda E|$ se llama *polinomio característico* de la matriz A, y sus raíces, que pueden ser tanto reales como imaginarias, se llaman *raíces características* de la misma.

Las matrices semejantes poseen iguales polinomios característicos y, por consiguiente, iguales raíces características.

En efecto, supongamos que

$$B=Q^{-1}AQ.$$

Entonces, teniendo en cuenta que la matriz λE es conmutable con la matriz Q y que $|Q^{-1}|=|Q|^{-1}$, *resulta:*

$$|B-\lambda E|=|Q^{-1}AQ-\lambda E|=|Q^{-1}(A-\lambda E)Q|=$$
$$=|Q|^{-1}\cdot|A-\lambda E|\cdot|Q|=|A-\lambda E|,$$

que es lo que se quería demostrar.

En virtud del teorema demostrado en el § 31 sobre la relación entre las matrices que determinan una transformación lineal en diferentes bases, se deduce que *a pesar de que la transformación lineal φ se puede determinar por diferentes matrices en bases diversas, sin embargo, estas matrices tienen un mismo conjunto de raíces características.* Por consiguiente, estas raíces se pueden llamar *raíces características de la misma transformación φ.* Todo el conjunto de estas raíces características, tomando cada raíz con el mismo orden

de multiplicidad que tiene en el polinomio característico, se llama *espectro* de la transformación lineal φ.

Las raíces características desempeñan un gran papel en el estudio de las transformaciones lineales. El lector tendrá muchas oportunidades para convencerse de esto. Por ahora, indicaremos una de las aplicaciones de las raíces características.

Supongamos que en el espacio lineal real V_n se ha dado una transformación lineal φ. Si en este caso el vector b, diferente de cero, se transforma en otro que es proporcional al mismo vector b,

$$b\varphi = \lambda_0 b, \tag{1}$$

donde λ_0 es cierto número real, el vector b se llama *vector propio* de la transformación φ y el número λ_0, *valor propio* de la misma; además, se dice que el vector propio b *corresponde* al valor propio λ_0.

Obsérvese que como $b \neq 0$, el número λ_0 que satisface a la condición (1) se determina para el vector b unívocamente. Subrayemos luego que el vector nulo no se considera vector propio de la transformación φ, a pesar de que satisface a la condición (1) para cualquier λ_0.

La rotación del plano euclídeo alrededor del origen de coordenadas, en un ángulo que no sea múltiplo de π es un ejemplo de transformación lineal que carece de vectores propios. Otro ejemplo de caso extremo es la dilatación del plano; aquí todos los vectores que parten del origen de coordenadas se alargan, aumentando de longitud por ejemplo, cinco veces. Para esta transformación lineal, son propios todos los vectores no nulos del plano; todos ellos corresponden al valor propio 5.

Las raíces características reales de la transformación lineal φ, (si es que éstas existen), y sólo éstas, son valores propios de la misma.

En efecto, supongamos que en la base $e_1, e_2, \ldots, e_n$, la transformación φ tiene la matriz $A = (\alpha_{ij})$ y sea

$$b = \sum_{i=1}^{n} \beta_i e_i$$

un vector propio de la transformación φ,

$$b\varphi = \lambda_0 b.$$

Como se ha demostrado en el § 31,

$$b\varphi = [(\beta_1, \beta_2, \ldots, \beta_n) A] e. \tag{3}$$

Las igualdades (2) y (3) dan lugar al sistema de igualdades

$$\left.\begin{aligned} \beta_1\alpha_{11} + \beta_2\alpha_{21} + \ldots + \beta_n\alpha_{n1} &= \lambda_0\beta_1, \\ \beta_1\alpha_{12} + \beta_2\alpha_{22} + \ldots + \beta_n\alpha_{n2} &= \lambda_0\beta_2, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ \beta_1\alpha_{1n} + \beta_2\alpha_{2n} + \ldots + \beta_n\alpha_{nn} &= \lambda_0\beta_n. \end{aligned}\right\} \tag{4}$$

Como $b \neq 0$, no todos los números $\beta_1, \beta_2, \ldots, \beta_n$ son iguales a *cero*. Por lo tanto, en virtud de (4), *el sistema de ecuaciones* lineales homogéneas

$$\begin{aligned} (\alpha_{11}-\lambda_0)\,x_1+\alpha_{21}x_2+\ldots+\alpha_{n1}x_n&=0,\\ \alpha_{12}x_1+(\alpha_{22}-\lambda_0)\,x_2+\ldots+\alpha_{n2}x_n&=0,\\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots&\\ \alpha_{1n}x_1+\alpha_{2n}x_2+\ldots+(\alpha_{nn}-\lambda_0)\,x_n&=0 \end{aligned} \tag{5}$$

posee solución no nula; de donde su determinante es igual a cero,

$$\begin{vmatrix} \alpha_{11}-\lambda_0, & \alpha_{21}, & \ldots, & \alpha_{n1}\\ \alpha_{12}, & \alpha_{22}-\lambda_0, & \ldots, & \alpha_{n2}\\ \ldots & \ldots & \ldots & \ldots\\ \alpha_{1n}, & \alpha_{2n}, & \ldots, & \alpha_{nn}-\lambda_0 \end{vmatrix}=0. \tag{6}$$

Transponiendo, resulta

$$|A-\lambda_0E|=0, \tag{7}$$

o sea, el valor propio λ_0 es, en efecto, raíz característica de la matriz A y, por consiguiente, de la transformación φ; naturalmente, éste es *real*.

Recíprocamente, supongamos que λ_0 es una raíz característica real cualquiera de la transformación φ y, por consiguiente, de la matriz A. Entonces, se cumple la igualdad (7) y, por lo tanto, también la igualdad (6) obtenida de la (7) por transposición. De aquí se deduce que el sistema de ecuaciones lineales homogéneas (5) posee solución no nula, e incluso real, puesto que todos los coeficientes de este sistema son reales. Indicando esta solución mediante

$$(\beta_1, \beta_2, \ldots, \beta_n), \tag{8}$$

se verifican las igualdades (4). Designemos con b el vector del espacio V_n que tiene la fila de coordenadas (8) en la base $e_1, e_2, \ldots, e_n$; está claro que $b \neq 0$. Entonces, se verifica la igualdad (3), y de (4) y (3) se deduce (2). *Por lo tanto, resulta que b es un vector propio* de la transformación φ, correspondiente al valor propio λ_0. El teorema queda demostrado.

Obsérvese que si se examinase un espacio lineal **complejo**, no habría necesidad de que la raíz característica fuese real, o sea, quedaría demostrado el teorema siguiente: *las raíces características de una transformación lineal del espacio lineal complejo, y sólo ellas, son valores propios de esta transformación*. De aquí se deduce que *toda transformación lineal posee vectores propios en el espacio lineal complejo*.

Volviendo al caso real considerado, señalemos que el conjunto de los vectores propios de la transformación lineal φ que corresponden al valor propio λ_0, coincide con el conjunto de soluciones reales no nulas del sistema de ecuaciones lineales homogéneas (5). De aquí se deduce que *el conjunto de vectores propios de la transformación lineal φ, correspondientes al valor propio λ_0, después de agregarle el vector nulo, es un subespacio lineal del espacio V_n.* En efecto, de lo demostrado en el § 12 se deduce que *el conjunto de las soluciones (reales) de cualquier sistema de ecuaciones lineales homogéneas con n incógnitas es un subespacio lineal del espacio V_n.*

Transformaciones lineales con espectro simple. En muchos casos se necesita saber si una transformación lineal dada φ puede tener en cierta base una matriz diagonal. En realidad, no cualquier transformación lineal se puede determinar por una matriz diagonal. En el § 61 se indicarán las condiciones necesarias y suficientes para esto; ahora queremos exponer una condición suficiente.

Demostremos primero las siguientes proposiciones auxiliares: *Una transformación lineal φ se determina en la base $e_1, e_2, \ldots, e_n$ por una matriz diagonal cuando, y sólo cuando, todos los vectores de esta base son vectores propios de la transformación φ.*

En efecto, la igualdad

$$e_i\varphi = \lambda_i e_i$$

equivale a que en la i-ésima fila de la matriz que determina la transformación lineal en la base indicada, sean iguales a cero todos los elementos que estén fuera de la diagonal principal, y que en la diagonal principal (o sea, en el i-ésimo lugar) figure el número λ_i.

Los vectores propios $b_1, b_2, \ldots, b_k$ de la transformación lineal φ que corresponden a diferentes valores propios, forman un sistema linealmente independiente.

Demostraremos esta afirmación por el método de inducción sobre k, puesto que para $k = 1$ se cumple: un vector propio, diferente de cero, forma un sistema linealmente independiente. Supongamos que

$$b_i\varphi = \lambda_i b_i, \quad i = 1, 2, \ldots, k,$$

donde

$$\lambda_i \neq \lambda_j \text{ para } i \neq j.$$

Si existiese una dependencia lineal

$$\alpha_1 b_1 + \alpha_2 b_2 + \ldots + \alpha_k b_k = 0, \tag{9}$$

donde, por ejemplo, $\alpha_1 \neq 0$, entonces aplicando a ambos miembros de la igualdad (9) la transformación φ, obtendríamos

$$\alpha_1\lambda_1 b_1 + \alpha_2\lambda_2 b_2 + \ldots + \alpha_k\lambda_k b_k = 0.$$

Restando de aquí la igualdad (9) multiplicada por λ_k, resultaría

$$\alpha_1(\lambda_1-\lambda_k)b_1+\alpha_2(\lambda_2-\lambda_k)b_2+\ldots+\alpha_{k-1}(\lambda_{k-1}-\lambda_k)b_{k-1}=0,$$

lo que da una dependencia lineal no trivial entre los vectores $b_1, b_2, \ldots, b_{k-1}$, pues $\alpha_1(\lambda_1-\lambda_k)\neq 0$.

Se dice que una transformación lineal φ del espacio lineal real V_n tiene un *espectro simple*, si todas sus raíces características son reales y distintas. Por consiguiente, la transformación φ tiene n valores propios distintos, de aquí que, en virtud del teorema demostrado, en el espacio V_n existe una base compuesta de vectores propios de esta transformación. Por lo tanto, *toda transformación lineal con espectro simple puede determinarse por una matriz diagonal.*

Pasando de la transformación lineal a las matrices que la determinan, obtenemos el resultado siguiente:

Toda matriz cuyas raíces características son reales y distintas, es semejante a una matriz diagonal o, como suele decirse, se reduce a la forma diagonal.

CAPITULO VIII

ESPACIOS EUCLIDEOS

§ 34. Definición del espacio euclídeo. Bases ortonormales

El concepto de espacio lineal de n dimensiones no generaliza en gran medida el concepto de plano euclídeo o de espacio euclídeo de tres dimensiones, pues, en el caso de n dimensiones, para $n > 3$ no está definida ni la longitud de un vector, ni el ángulo entre los vectores, resultando imposible el desarrollo de la rica teoría geométrica, que conoce bien el lector para $n = 2$ y $n = 3$. Sin embargo la situación tiene salida del modo siguiente.

Por el curso de geometría analítica se sabe que en el plano y en el espacio de tres dimensiones se puede introducir el concepto de producto escalar de vectores. Este se define mediante la longitud de los vectores y el ángulo formado por ellos. No obstante, resulta que la longitud del vector y el ángulo formado por los vectores se pueden expresar a su vez mediante los productos escalares. Por esto, definiremos axiomáticamente el producto escalar en cualquier espacio lineal de n dimensiones mediante algunas propiedades bien conocidas del producto escalar de vectores en el plano o en el espacio de tres dimensiones. Además, teniendo en cuenta los objetivos inmediatos, debido a los cuales fue introducido este apartado en el curso de álgebra superior, aquí no se dará la definición de longitud de un vector y de ángulo entre los vectores. Al lector que le interese la construcción de la geometría en el espacio de n dimensiones le recomendamos que consulte literatura más especializada; en primer lugar, la del álgebra lineal.

En este capítulo, a excepción del final del presente párrafo, se examinan siempre los espacios lineales **reales.**

Diremos que en el espacio lineal real V_n de n dimensiones está definido el *producto escalar*, si a cada par de vectores a, b se ha puesto en correspondencia un número real, designado por la notación (a, b) y denominado *producto escalar* de los vectores a y b, cumpliéndose las condiciones siguientes (aquí, a, b, c son vectores arbitrarios del espacio V_n, α es un número real cualquiera):

I. $$(a, b) = (b, a).$$

II. $(a+b, c)=(a, c)+(b, c)$.

III. $(\alpha a, b)=\alpha(a, b)$.

IV. Si $a \neq 0$, el cuadrado escalar del vector a es estrictamente positivo,

$$(a, a)>0.$$

De III, para $\alpha=0$ se deduce la igualdad

$$(0, b)=0, \tag{1}$$

o sea, *el producto escalar del vector nulo por cualquier vector b es igual a cero*; en particular, es igual a cero el cuadrado escalar del vector nulo.

De II y III, se obtiene inmediatamente la siguiente **fórmula para el producto escalar de las combinaciones lineales de dos sistemas de vectores:**

$$\left(\sum_{i=1}^{k} \alpha_i a_i, \sum_{j=1}^{l} \beta_j b_j\right)=\sum_{i=1}^{k} \alpha_i \beta_j (a_i, b_j). \tag{2}$$

Si en el espacio lineal de n dimensiones está definido el producto escalar, éste se llama *espacio euclídeo* de n dimensiones.

Para cualquier n, se puede definir el producto escalar en el espacio lineal V_n de n dimensiones, o sea, este espacio se puede convertir en euclídeo.

En efecto, tomemos en el espacio V_n cualquier base $e_1, e_2, \ldots, e_n$. Si

$$a=\sum_{i=1}^{n} \alpha_i e_i, \quad b=\sum_{i=1}^{n} \beta_i e_i,$$

suponemos,

$$(a, b)=\sum_{i=1}^{n} \alpha_i \beta_i. \tag{3}$$

Fácilmente se comprueba que se cumplen las condiciones I-IV, o sea, que la igualdad (1) determina el producto escalar en el espacio V_n.

Vemos, pues, que generalmente, en el espacio lineal de n dimensiones se puede definir el producto escalar de muchos modos; naturalmente, la definición (3) depende de la elección de la base. Sin embargo, no sabemos por ahora si el producto escalar se puede introducir de algún modo que de principio sea diferente. Nuestro próximo objetivo consiste en examinar todos los modos posibles de convertir el espacio lineal de n dimensiones en espacio euclídeo y en establecer que, en cierto sentido, para cualquier n existe un solo espacio euclídeo de n dimensiones.

Sea dado un espacio euclídeo arbitrario E_n de n dimensiones, o sea, que en el espacio lineal de n dimensiones está definido arbitrariamente el producto escalar. Se dice que los vectores a y b son *ortogonales*, si su producto escalar es igual a cero,

$$(a, b) = 0.$$

De (1) se deduce que el vector nulo es ortogonal a cualquier vector; sin embargo, pueden existir también vectores ortogonales no nulos.

Un sistema de vectores se denomina *sistema ortogonal*, si todos los vectores de este sistema son ortogonales entre sí dos a dos.

Todo sistema ortogonal de vectores no nulos es linealmente independiente.

En efecto, sea dado en E_n un sistema de vectores $a_1, a_2, \ldots, a_k$, donde $a_i \neq 0$, $i = 1, 2, \ldots, k$, y

$$(a_i, a_j) = 0 \text{ para } i \neq j. \tag{4}$$

Si

$$\alpha_1 a_1 + \alpha_2 a_2 + \ldots + \alpha_k a_k = 0,$$

multiplicando escalarmente ambos miembros de esta igualdad por el vector a_i, $1 \leqslant i \leqslant k$, en virtud de (1), (2) y (4), resulta:

$$0 = (0, a_i) = (\alpha_1 a_1 + \alpha_2 a_2 + \ldots + \alpha_k a_k, a_i) =$$
$$= \alpha_1 (a_1, a_i) + \alpha_2 (a_2, a_i) + \ldots + \alpha_k (a_k, a_i) = \alpha_i (a_i, a_i).$$

De aquí, como $(a_i, a_i) > 0$ según IV, se tiene $\alpha_i = 0$, $i = 1, 2, \ldots$ $\ldots, k$, como se quería demostrar.

Ahora se va a describir el **proceso de ortogonalización**, o sea, un método para pasar de cualquier sistema de k vectores

$$a_1, a_2, \ldots, a_k \tag{5}$$

linealmente independiente, del espacio euclídeo E_n, a un sistema ortogonal, compuesto también de k vectores no nulos; estos vectores se indicarán mediante $b_1, b_2, \ldots, b_k$.

Hagamos $b_1 = a_1$, de modo que **el primer vector del sistema** (5) **quedará incluido en el sistema ortogonal que se construye.** Hagamos luego

$$b_2 = \alpha_1 b_1 + a_2.$$

Como $b_1 = a_1$ y los vectores a_1 y a_2 son linealmente independientes, el vector b_2 será diferente de cero para cualquier número α_1. Elijamos este número de modo que el vector b_2 sea ortogonal al vector b_1:

$$0 = (b_1, b_2) = (b_1, \alpha_1 b_1 + a_2) = \alpha_1 (b_1, b_1) + (b_1, a_2),$$

de donde, en virtud de IV,

$$\alpha_1 = -\frac{(b_1, a_2)}{(b_1, b_1)}.$$

Supongamos que ya está construido el sistema ortogonal de vectores no nulos $b_1, b_2, \ldots, b_l$; supongamos, además, que para *cualquier* i, $1 \leqslant i \leqslant l$, *el vector* b_i *es combinación lineal de los vectores* $a_1, a_2, \ldots, a_i$. Entonces, esto mismo se cumplirá también para el vector b_{l+1}, si éste se elige de la forma siguiente:

$$b_{l+1} = \alpha_1 b_1 + \alpha_2 b_2 + \ldots + \alpha_l b_l + a_{l+1}.$$

En este caso, el vector b_{l+1} será diferente de cero, puesto que el sistema (5) es linealmente independiente y el vector a_{l+1} no está incluido entre los vectores $b_1, b_2, \ldots, b_l$. Los coeficientes α_i, $i = 1, 2, \ldots, l$ se eligen de modo que el vector b_{l+1} sea ortogonal a todos los vectores b_i, $i = 1, 2, \ldots, l$:

$$0 = (b_i, b_{l+1}) = (b_i, \alpha_1 b_1 + \alpha_2 b_2 + \ldots + \alpha_l b_l + a_{l+1}) =$$
$$= \alpha_1 (b_i, b_1) + \alpha_2 (b_i, b_2) + \ldots + \alpha_l (b_i, b_l) + (b_i, a_{l+1});$$

de aquí, como los vectores $b_1, b_2, \ldots, b_l$ son ortogonales entre sí, resulta

$$\alpha_i (b_i, b_i) + (b_i, a_{l+1}) = 0,$$

o sea,

$$\alpha_i = -\frac{(b_i, a_{l+1})}{(b_i, b_i)}, \quad i = 1, 2, \ldots, l.$$

Continuando este proceso se construye el sistema ortogonal buscado $b_1, b_2, \ldots, b_k$.

Aplicando este proceso de ortogonalización a cualquier base del espacio E_n, se obtiene un sistema ortogonal de n vectores no nulos, es decir una *base ortogonal*, pues, por lo demostrado, este sistema es linealmente independiente. Recordando ahora la observación hecha con relación al primer paso del proceso de ortogonalización y teniendo en cuenta también que cualquier vector no nulo se puede incluir en una base del espacio, se puede enunciar incluso la siguiente afirmación:

Todo espacio euclídeo posee bases ortogonales; cualquier vector no nulo de este espacio forma parte de alguna base ortogonal.

En adelante, desempeñará un papel importante una forma especial de base ortogonal; las bases de esta forma corresponden a los sistemas cartesianos rectangulares de coordenadas empleados en la geometría analítica.

El vector b se llamará *normal*, si su cuadrado escalar es igual a la unidad,

$$(b, b) = 1.$$

Si $a \neq 0$, de donde $(a, a) > 0$, el paso al vector

$$b = \frac{1}{\sqrt{(a, a)}} a$$

se denominará *normalización* del vector a. El vector b es normal, puesto que

$$(b, b)=\left(\frac{1}{\sqrt{(a, a)}}a, \frac{1}{\sqrt{(a, a)}}a\right)=\left(\frac{1}{\sqrt{(a, a)}}\right)^2 (a, a)=1.$$

Una base $e_1, e_2, \ldots, e_n$ del espacio euclídeo E_n se llama *ortonormal*, si ésta es ortogonal y todos sus vectores son normales, es decir, si

$$\begin{aligned}(e_i, e_j)&=0 \text{ para } i\neq j,\\ (e_i, e_i)&=1, \qquad i=1, 2, \ldots, n.\end{aligned}\tag{6}$$

Todo espacio euclídeo posee bases ortonormales.

Para la demostración es suficiente tomar cualquier base ortogonal y normalizar todos sus vectores. Con esto, la base se mantiene ortogonal, puesto que, para cualesquiera α y β, de $(a, b)=0$, se deduce que

$$(\alpha a, \beta b)=\alpha\beta (a, b)=0.$$

La base $e_1, e_2, \ldots, e_n$ del espacio euclídeo E_n es ortonormal si, y sólo si, el producto escalar de dos vectores cualesquiera del espacio es igual a la suma de los productos de las coordenadas correspondientes de estos vectores en la base indicada, o sea, si de

$$a=\sum_{i=1}^{n}\alpha_i e_i, \quad b=\sum_{j=1}^{n}\beta_j e_j \tag{7}$$

se deduce que

$$(a, b)=\sum_{i=1}^{n}\alpha_i\beta_i. \tag{8}$$

En efecto, si para nuestra base se verifican las igualdades (6), se tiene

$$(a, b)=\left(\sum_{i=1}^{n}\alpha_i e_i, \sum_{j=1}^{n}\beta_j e_j\right)=\sum_{i, j=1}^{n}\alpha_i\beta_j (e_i, e_j)=\sum_{i=1}^{n}\alpha_i\beta_i.$$

Recíprocamente, si nuestra base es tal, que para cualesquiera vectores a y b expresados en esta base en la forma (7) se cumple la igualdad (8), entonces, tomando por a y b dos vectores cualesquiera de esta base e_i y e_j, iguales o diferentes, las igualdades (6) se obtendrán de (8).

Confrontando este resultado obtenido con la demostración anterior de la existencia de espacios euclídeos de n dimensiones para cualquier n, se puede enunciar la siguiente proposición: *si en el espacio lineal de n dimensiones V_n se ha elegido una base arbitraria, en V_n se puede definir el producto escalar de modo que en el espacio euclídeo obtenido la base elegida sea una de las bases ortonormales.*

Isomorfismo de los espacios euclídeos. Se dice que los espacios euclídeos E y E' son *isomorfos*, si entre los vectores de los mismos se puede establecer una correspondencia biunívoca tal, que se cumplan las condiciones siguientes:

1) esta correspondencia es una correspondencia de isomorfismo entre E y E', considerados éstos como espacios lineales (véase el § 29);

2) *en esta correspondencia se conserva el producto escalar; en* otras palabras, si las imágenes de los vectores a y b de E son los vectores a' y b' de E', respectivamente, entonces

$$(a, b) = (a', b'). \tag{9}$$

De la condición 1) se deduce inmediatamente que *los espacios euclídeos isomorfos tienen una misma dimensión.* Demostremos la afirmación recíproca:

Dos espacios euclídeos cualesquiera E y E' que tengan una dimensión n, son isomorfos entre sí.

En efecto, elijamos en los espacios E y E' las bases **ortonormales**

$$e_1, e_2, \ldots, e_n \tag{10}$$

y, respectivamente,

$$e'_1, e'_2, \ldots, e'_n. \tag{11}$$

Poniendo en correspondencia a cada vector

$$a = \sum_{i=1}^{n} \alpha_i e_i$$

de E el vector

$$a' = \sum_{i=1}^{n} \alpha_i e'_i$$

de E', que tiene en la base (11) las mismas coordenadas que tiene el vector a en la base (10), se obtiene, evidentemente, una correspondencia de isomorfismo entre los espacios **lineales** E y E'. Demostremos que también se cumple la igualdad (9): si

$$b = \sum_{i=1}^{n} \beta_i e_i, \quad b' = \sum_{i=1}^{n} \beta_i e'_i,$$

entonces, en virtud de (8) (téngase en cuenta que las bases (10) y (11) son ortonormales),

$$(a, b) = \sum_{i=1}^{n} \alpha_i \beta_i = (a', b').$$

Es natural que los espacios euclídeos isomorfos no se deben considerar diferentes. Por esto, para cualquier n, existe solamente un espacio euclídeo de n dimensiones en el mismo sentido en que para cualquier n existe solamente un espacio lineal real de n dimensiones.

Para el caso de espacios lineales **complejos,** los conceptos y resultados del presente párrafo se generalizan del modo siguiente: un espacio lineal complejo se llama *espacio unitario*, si en él está definido el producto escalar, donde (a, b) es, por lo general, un número complejo. En este caso, tienen que cumplirse los axiomas II—IV (en el enunciado del último axioma se debe subrayar que el cuadrado escalar de todo vector no nulo es real y estrictamente positivo); el axioma I tiene que sustituirse por el axioma

I' $$(a, b)=\overline{(b, a)},$$

donde la raya señala, como es usual, el paso al número complejo conjugado.

Por consiguiente, el producto escalar ya no es conmutativo. A pesar de todo, se verifica una igualdad simétrica al axioma II,

II' $$(a, b+c)=(a, b)+(a, c),$$

ya que

$$(a, b+c)=\overline{(b+c, a)}=\overline{(b, a)+(c, a)}=\overline{(b, a)}+\overline{(c, a)}=(a, b)+(a, c).$$

Por otra parte,

III' $$(a, \alpha b)=\bar{\alpha}(a, b),$$

puesto que

$$(a, \alpha b)=\overline{(\alpha b, a)}=\overline{\alpha(b, a)}=\bar{\alpha}\,\overline{(b, a)}=\bar{\alpha}(a, b).$$

Los conceptos de sistema ortogonal y ortonormal de vectores se generalizan sin alteración alguna al caso de espacios unitarios. Igual que antes, se demuestra la existencia de bases ortonormales en cualquier espacio unitario de dimensión finita. Sin embargo, si en este caso $e_1, e_2, \ldots, e_n$ es una base ortonormal y los vectores a, b tienen en esta base la expresión (7), se tiene

$$(a, b)=\sum_{i=1}^{n} \alpha_i\bar{\beta}_i.$$

Los resultados de los párrafos posteriores del presente capítulo también se podrían generalizar para el caso de espacios unitarios. Aquí no lo haremos y proponemos al lector que le interese que consulte libros especializados en álgebra lineal.

§ 35. Matrices ortogonales, transformaciones ortogonales

Sea dada una transformación lineal real de n indeterminadas

$$x_i=\sum_{k=1}^{n} q_{ik}y_k, \quad i=1, 2, \ldots, n; \tag{1}$$

la matriz de esta transformación se denotará por Q. Esta transformación lleva la suma de cuadrados de las indeterminadas $x_1, x_2, \ldots, x_n$, o sea, la forma cuadrática $x_1^2+x_2^2+\ldots+x_n^2$, que es la forma normal de las formas cuadráticas definidas positivas (véase el § 28), a cierta forma en las indeterminadas $y_1, y_2, \ldots, y_n$. Eventualmente, esta nueva forma cuadrática también puede resultar ser la suma de los cuadrados de las indeterminadas $y_1, y_2, \ldots, y_n$,

es decir, puede ocurrir que se verifique la igualdad

$$x_1^2+x_2^2+\ldots+x_n^2=y_1^2+y_2^2+\ldots+y_n^2, \qquad (2)$$

lo cual se convierte en una identidad después de sustituir las indeterminadas $x_1, x_2, \ldots, x_n$ por sus expresiones (1). La transformación lineal de las indeterminadas (1) que posee esta propiedad, o como suele decirse, que mantiene invariante la suma de los cuadrados de las indeterminadas, se llama *transformación ortogonal de las indeterminadas*, y su matriz Q, *matriz ortogonal*.

Existen muchas otras definiciones de transformación ortogonal y de matriz ortogonal, equivalentes a la expuesta anteriormente. Indiquemos algunas de ellas que emplearemos después.

Ya conocemos, por el § 26, la ley según la cual se transforma la matriz de una forma cuadrática al realizar una transformación lineal de las indeterminadas. Aplicándola a nuestro caso y teniendo en cuenta que la matriz de la forma cuadrática, que es la suma *de los cuadrados de todas las indeterminadas, es la matriz unidad* E, resulta que la igualdad (2) es equivalente a la igualdad matricial

$$Q'EQ=E,$$

o sea,

$$Q'Q=E. \qquad (3)$$

De aquí que

$$Q'=Q^{-1}, \qquad (4)$$

por lo que también se cumple la igualdad

$$QQ'=E. \qquad (5)$$

Por consiguiente, en virtud de (4), *la matriz ortogonal Q se puede definir como una matriz para la que la matriz transpuesta Q' es igual a la matriz inversa Q^{-1}*. Cada una de las igualdades (3) y (5) se puede tomar también por definición de matriz ortogonal.

Como las columnas de la matriz Q' son filas de la matriz Q, de (5) se deduce la proposición siguiente: *la matriz cuadrada Q es ortogonal cuando, y sólo cuando, la suma de los cuadrados de todos los elementos de cualquiera de sus filas es igual a la unidad y la suma de los productos de los elementos correspondientes de dos filas cualesquiera diferentes es igual a cero*. De (3) se deduce la proposición análoga para las columnas de la matriz Q.

Como $Q' = Q$, pasando a determinantes en la igualdad (3), resulta la igualdad

$$|Q|^2=1.$$

De aquí se deduce que *el determinante de una matriz ortogonal es igual a* ± 1. Por lo tanto, *toda transformación ortogonal de las inde-*

terminadas es no degenerada. Naturalmente, no se puede afirmar lo recíproco; señalemos también que no cualquier matriz con el determinante igual a ± 1 es ortogonal.

La matriz inversa de una matriz ortogonal es también ortogonal. En efecto, pasando en (4) a las matrices traspuestas, resulta:

$$(Q^{-1}) = (Q')' = Q = (Q^{-1})^{-1}.$$

Por otra parte, *el producto de matrices ortogonales también es ortogonal*. En efecto, si las matrices Q y R son ortogonales, entones aplicando (4) y también la igualdad (6) del § 26 y la igualdad análoga que se verifica para la matriz inversa, resulta:

$$(QR)' = R'Q' = R^{-1}Q^{-1} = (QR)^{-1}.$$

En el § 37 se empleará la proposición siguiente:

La matriz de cambio para pasar de una base ortonormal del espacio euclídeo a otra base ortonormal cualquiera es ortogonal.

En efecto, supongamos que en el espacio E_n se han dado dos bases ortonormales $e_1, e_2, \ldots, e_n$ y $e'_1, e'_2, \ldots, e'_n$ con la matriz de cambio $Q = (q_{ij})$,

$$e' = Qe.$$

Como la base e es ortonormal, el producto escalar de dos vectores cualesquiera y, en particular, de dos vectores cualesquiera de la base e', es igual a la suma de los productos de las coordenadas correspondientes de estos vectores en la base e. Sin embargo, como la base e' también es ortonormal, el cuadrado escalar de cada vector de e' es igual a la unidad, y el producto escalar de dos vectores diversos cualesquiera de e' es igual a cero. De aquí, para las filas de las coordenadas de los vectores de la base e' en la base e, o sea, para las filas de la matriz Q, resultan las afirmaciones que son características para una matriz ortogonal, como se dedujo antes de la igualdad (5).

Transformaciones ortogonales del espacio euclídeo. Ahora se estudiará un tipo especial e interesante de transformaciones lineales de los espacios euclídeos, a pesar de que estas transformaciones no se emplearán a continuación.

Una transformación lineal φ del espacio euclídeo E_n se llama *transformación ortogonal* de éste, si mantiene invariable el cuadrado escalar de cada vector, es decir, si para cualquier vector a

$$(a\varphi, a\varphi) = (a, a). \tag{6}$$

Ahora deduciremos la siguiente proposición más general que la anterior y que, naturalmente, también se puede tomar por definición de transformación ortogonal.

Toda transformación ortogonal φ *del espacio euclídeo mantiene invariable el producto escalar de dos vectores cualesquiera* a *y* b,

$$(a\varphi,\ b\varphi) = (a,\ b). \tag{7}$$

En efecto, en virtud de (6),

$$((a+b)\varphi,\ (a+b)\varphi) = (a+b,\ a+b).$$

Pero

$$((a+b)\varphi,\ (a+b)\varphi) = (a\varphi + b\varphi,\ a\varphi + b\varphi) =$$
$$= (a\varphi,\ a\varphi) + (a\varphi,\ b\varphi) + (b\varphi,\ a\varphi) + (b\varphi,\ b\varphi),$$
$$(a+b,\ a+b) = (a,\ a) + (a,\ b) + (b,\ a) + (b,\ b).$$

De aquí, aplicando (6) para a y para b, y teniendo en cuenta la propiedad conmutativa del producto escalar, resulta

$$2\,(a\varphi,\ b\varphi) = 2\,(a,\ b),$$

de donde, también se verifica (7).

En una transformación ortogonal del espacio euclídeo, las imágenes de todos los vectores de cualquier base ortonormal forman ellas mismas una base ortonormal. Recíprocamente, si una transformación lineal del espacio euclídeo transforma por lo menos una base ortonormal en otra base ortonormal, esta transformación es ortogonal.

En efecto, sea φ una transformación ortogonal del espacio sea E_n, y sea $e_1, e_2, \ldots, e_n$ una base ortonormal arbitraria de este espacio. En virtud de (7), de las igualdades

$$(e_i,\ e_i) = 1,\ i = 1,\ 2,\ \ldots,\ n,$$
$$(e_i,\ e_j) = 0 \text{ para } i \neq j$$

se deducen las igualdades

$$(e_i\varphi,\ e_i\varphi) = 1, \quad i = 1,\ 2,\ \ldots,\ n,$$
$$(e_i\varphi,\ e_j\varphi) = 0 \quad \text{para } i \neq j,$$

o sea, el sistema de vectores $e_1\varphi, e_2\varphi, \ldots, e_n\varphi$ resulta ortogonal y normal, por lo cual, éste es una base ortonormal del espacio E_n.

Recíprocamente, supongamos que la transformación lineal φ del espacio E_n transforma la base ortonormal $e_1, e_2, \ldots, e_n$ de nuevo en una base ortonormal, es decir, que el sistema de vectores $e_1\varphi, e_2\varphi, \ldots, e_n\varphi$ es una base ortonormal del espacio E_n. Si

$$a = \sum_{i=1}^{n} \alpha_i e_i$$

es un vector arbitrario del espacio E_n, entonces

$$a\varphi = \sum_{i=1}^{n} \alpha_i\,(e_i\varphi),$$

o sea, el vector $a\varphi$ tiene en la base $e\varphi$ las mismas coordenadas que tiene el vector a en la base e. Sin embargo, ambas bases son ortonormales, siendo, por consiguiente, el cuadrado escalar de cualquier vector igual a la suma de los cuadrados de sus coordenadas en cualquiera de estas bases. Por lo tanto,

$$(a, a) = (a\varphi, a\varphi) = \sum_{i=1}^{n} \alpha_i^2,$$

o sea, se cumple verdaderamente la igualdad (6).

Toda transformación ortogonal del espacio euclídeo se determina en cualquier base ortonormal por una matriz ortogonal. Recíprocamente, si una transformación lineal del espacio euclídeo se determina por una matriz ortogonal, aunque sólo sea en una base ortonormal, esta transformación es ortogonal.

En efecto, si la transformación φ es ortogonal y la base $e_1, e_2, \ldots, e_n$ es ortonormal, el sistema de vectores $e_1\varphi, e_2\varphi, \ldots, e_n\varphi$ será una base ortonormal. Por consiguiente, la matriz A de la transformación φ en la base e,

$$e\varphi = Ae, \tag{8}$$

será la matriz de cambio de la base ortonormal e por la base ortonormal $e\varphi$, y, por lo demostrado anteriormente, es ortogonal.

Recíprocamente, supongamos que la transformación lineal φ se determina en la base ortonormal $e_1, e_2, \ldots, e_n$ por la matriz ortogonal A; por consiguiente, se cumple la igualdad (8). Como la base e es ortonormal, el producto escalar de cualesquiera vectores y, en particular, de cualesquiera vectores del sistema $e_1\varphi, e_2\varphi, \ldots, e_n\varphi$, es igual a la suma de los productos de las coordenadas correspondientes de estos vectores en la base e. De donde, como la matriz A es ortogonal, se tiene

$$(e_i\varphi, e_i\varphi) = 1, \quad i = 1, 2, \ldots, n$$
$$(e_i\varphi, e_j\varphi) = 0 \quad \text{para } i \neq j,$$

o sea, resulta que el mismo sistema e es una base ortonormal del espacio E_n. De aquí se deduce que la transformación φ es ortogonal.

Por la geometría analítica sabemos, que entre todas las transformaciones afines del plano que dejan en su sitio el origen de coordenadas, las rotaciones (unidas, posiblemente, con simetrías) son las únicas que mantienen invariante el producto escalar. Por lo tanto, las transformaciones ortogonales del espacio euclídeo de n dimensiones se pueden considerar como «rotaciones» de este espacio.

Evidentemente, entre las transformaciones ortogonales del espacio euclídeo está también la transformación idéntica. Por otra parte, la relación que hemos establecido entre las transformaciones ortogonales y las matrices ortogonales, y también la relación expuesta

en el § 31 entre las operaciones con las transformaciones lineales y con las matrices, permiten deducir de las propiedades conocidas de las matrices ortogonales las siguientes propiedades de las transformaciones ortogonales del espacio euclídeo, que también se comprueban directamente con facilidad:

Toda transformación ortogonal es no degenerada y su transformación inversa también es ortogonal.

El producto de cualesquiera transformaciones ortogonales es ortogonal.

§ 36. Transformaciones simétricas

Una transformación lineal del espacio euclídeo de n dimensiones se llama *simétrica* (o bien, *autoconjugada*) si para cualesquiera vectores a, b de este espacio se verifica la igualdad

$$(a\varphi,\ b) = (a,\ b\varphi) \tag{1}$$

o sea, en el producto escalar el símbolo de la transformación simétrica se puede trasladar de un factor a otro.

Evidentemente, la transformación idéntica ε y la transformación nula ω son ejemplos de transformaciones simétricas. Un ejemplo más general es la transformación lineal, según la cual cada vector se multiplica por un número fijado α,

$$a\varphi = \alpha a.$$

En efecto, en este caso

$$(a\varphi,\ b) = (\alpha a,\ b) = \alpha (a,\ b) = (a,\ \alpha b) = (a,\ b\varphi).$$

Las transformaciones simétricas desempeñan un papel muy importante y es necesario estudiarlas detalladamente.

Toda transformación simétrica del espacio euclídeo se determina en cualquier base ortonormal por una matriz simétrica. Recíprocamente, si una transformación lineal del espacio euclídeo se determina por una matriz simétrica, aunque sólo sea en una base ortonormal, la transformación es simétrica.

En efecto, supongamos que la transformación simétrica φ se determina por la matriz $A = (\alpha_{ij})$ en la base ortonormal $e_1, e_2, \ldots$ $\ldots, e_n$. Teniendo en cuenta que en una base ortonormal, el producto escalar de dos vectores es igual a la suma de los productos de las coordenadas correspondientes de estos vectores, se obtiene:

$$(e_i\varphi,\ e_j) = \Big(\sum_{k=1}^{n} \alpha_{ik}e_k,\ e_j\Big) = \alpha_{ij},$$

$$(e_i,\ e_j\varphi) = \Big(e_i,\ \sum_{k=1}^{n} \alpha_{jk}e_k\Big) = \alpha_{ji},$$

o sea, en virtud de (1),

$$\alpha_{ij}=\alpha_{ji}$$

para todos los valores de i y j. De aquí, la matriz A es simétrica.

Recíprocamente, supongamos que la transformación lineal φ se determina por una matriz simétrica $A=(\alpha_{ij})$ en la base ortonormal $e_1, e_2, \ldots, e_n$,

$$\alpha_{ij}=\alpha_{ji} \text{ para todos los valores de } i \text{ y } j. \tag{2}$$

Si

$$b=\sum_{i=1}^{n}\beta_i e_i, \quad c=\sum_{j=1}^{n}\gamma_j e_j$$

son unos vectores arbitrarios del espacio, entonces

$$b\varphi=\sum_{i=1}^{n}\beta_i(e_i\varphi)=\sum_{j=1}^{n}\left(\sum_{i=1}^{n}\beta_i\alpha_{ij}\right)e_j,$$

$$c\varphi=\sum_{j=1}^{n}\gamma_j(e_j\varphi)=\sum_{i=1}^{n}\left(\sum_{j=1}^{n}\gamma_j\alpha_{ji}\right)e_i.$$

Teniendo en cuenta que la base e es ortonormal, resulta

$$(b\varphi, c)=\sum_{j,\, i=1}^{n}\beta_i\alpha_{ij}\gamma_j,$$

$$(b, c\varphi)=\sum_{i,\, j=1}^{n}\beta_i\gamma_j\alpha_{ji}.$$

En virtud de (2), los segundos miembros de las últimas igualdades coinciden, así que

$$(b\varphi, c)=(b, c\varphi),$$

que es lo que se quería demostrar.

Del resultado obtenido se deduce la siguiente propiedad de las transformaciones simétricas, que también se comprueba con facilidad directamente:

La suma de transformaciones simétricas y también el producto de una transformación simétrica por un número, son transformaciones simétricas.

Demostremos ahora el siguiente importante teorema:

Todas las raíces características de una transformación simétrica son reales.

Como las raíces características de cualquier transformación lineal coinciden con las raíces características de la matriz de esta transformación en cualquier base y la transformación simétrica se determina en bases ortonormales por matrices simétricas, es suficiente demostrar la proposición siguiente:

Todas las raíces características de una matriz simétrica son reales.

En efecto, sea λ_0 una raíz característica (que puede ser compleja) de la matriz simétrica $A = (\alpha_{ij})$,

$$|A - \lambda_0 E| = 0.$$

Entonces el sistema de ecuaciones lineales homogéneas con coeficientes *complejos*

$$\sum_{j=1}^{n} \alpha_{ij} x_j = \lambda_0 x_i, \quad i = 1, 2, \ldots, n,$$

tiene un determinante igual a cero, o sea, posee solución **no nula** $\beta_1, \beta_2, \ldots, \beta_n$, *que por lo general, es compleja*; por lo tanto,

$$\sum_{i=1}^{n} \alpha_{ij} \beta_j = \lambda_0 \beta_i, \ i = 1, 2, \ldots, n. \tag{3}$$

Multiplicando ambos miembros de cada i-ésima igualdad (3) por el número $\overline{\beta}_i$, conjugado con el número β_i, y sumando por separado los primeros y los segundos miembros de todas las igualdades obtenidas, se llega a la igualdad

$$\sum_{i,\, j=1}^{n} a_{ij} \beta_j \overline{\beta}_i = \lambda_0 \sum_{i=1}^{n} \beta_i \overline{\beta}_i. \tag{4}$$

El coeficiente de λ_0 en la igualdad (4) es un número real, diferente de cero, puesto que es la suma de números reales no negativos, uno de los cuales por lo menos es estrictamente positivo. Para demostrar *que el número λ_0 es real*, hay que demostrar que es real el primer miembro de la igualdad (4), para lo cual es suficiente demostrar que este número complejo coincide con su conjugado. Aquí, por primera vez se aplicará la simetría de la matriz (real) A:

$$\overline{\sum_{i,\, j=1}^{n} \alpha_{ij} \beta_j \overline{\beta}_i} = \sum_{i,\, j=1}^{n} \overline{\alpha_{ij} \beta_j \overline{\beta}_i} = \sum_{i,\, j=1}^{n} \alpha_{ij} \overline{\beta}_j \beta_i =$$
$$= \sum_{i,\, j=1}^{n} \alpha_{ji} \overline{\beta}_j \beta_i = \sum_{i,\, j=1}^{n} \alpha_{ij} \overline{\beta}_i \beta_j = \sum_{i,\, j=1}^{n} \alpha_{ij} \beta_j \overline{\beta}_i.$$

Obsérvese que la penúltima igualdad se ha obtenido mediante una permutación simple de las notaciones de los índices de la suma: en lugar de i se ha puesto j, y en lugar de j se ha puesto i. Por consiguiente, el teorema queda demostrado.

Una transformación lineal φ del espacio euclídeo E_n es simétrica cuando, y sólo cuando, en el espacio E_n existe una base ortonormal formada por vectores propios de esta transformación.

Una parte de esta proposición es casi evidente: si en E_n existe una base *ortonormal* $e_1, e_2, \ldots, e_n$ tal que

$$e_i \varphi = \lambda_i e_i, \ i = 1, 2, \ldots, n,$$

entonces en la base e la transformación φ se determina por la matriz diagonal

$$\begin{pmatrix} \lambda_1 & & & & 0 \\ & \lambda_2 & & & \\ & & \cdot & & \\ & & & \cdot & \\ 0 & & & \cdot & \lambda_n \end{pmatrix}.$$

Pero, como toda matriz diagonal es simétrica resulta que la transformación φ se determina en la base ortonormal e por una matriz simétrica, es decir, ella misma es simétrica.

La proposición recíproca fundamental se demostrará por inducción sobre la dimensión n del espacio E_n. En efecto, para $n = 1$ cualquier transformación lineal φ del espacio E_1 lleva cualquier vector a otro que es proporcional al mismo. De esto se deduce que todo vector a no nulo es un vector propio para φ (por cierto, resulta también que toda transformación lineal del espacio E_1 es simétrica). Normalizando el vector a se obtiene la base ortonormal buscada del espacio E_1.

Supongamos que la tesis del teorema ya está demostrada para los espacios euclídeos de $(n - 1)$ dimensiones y que en el espacio E_n se ha dado una transformación simétrica φ. Del teorema demostrado anteriormente se deduce la existencia de una raíz característica real λ_0 para φ. Por consiguiente, para la transformación φ, este número es un valor propio. Si a es el vector propio de la transformación φ que corresponde a este valor propio, entonces cualquier vector no nulo que sea proporcional al vector a será para φ un vector propio correspondiente al mismo valor propio λ_0, puesto que

$$(\alpha a)\,\varphi = \alpha\,(a\varphi) = \alpha\,(\lambda_0 a) = \lambda_0\,(\alpha a).$$

En particular, normalizando el vector a, resulta un vector e_1 tal que

$$e_1\varphi = \lambda_0 e_1,$$
$$(e_1,\ e_1) = 1.$$

Como se demostró en el § 34, el vector no nulo e_1 se puede incluir en una base ortogonal

$$e_1,\ e_2',\ \ldots,\ e_n' \tag{5}$$

del espacio E_n. Los vectores cuyas primeras coordenadas en la base (5) son iguales a cero, o sea, los vectores de la forma $\alpha_2 e_2' + \ldots + \alpha_n e_n'$, forman, evidentemente, un subespacio lineal de $(n - 1)$ dimensiones del espacio E_n, que lo designaremos mediante L. Este será, incluso, un espacio euclídeo de $(n - 1)$ dimensiones, pues, estando definido

el producto escalar para todos los vectores de E_n, lo estará particularmente para los vectores de L, poseyendo además todas las propiedades necesarias.

El subespacio L consta de todos los vectores del espacio E_n que son ortogonales al vector e_1. En efecto, si

$$a = \alpha_1 e_1 + \alpha_2' e_2' + \ldots + \alpha_n' e_n',$$

como la base (5) es ortogonal y el vector e_1, normal, resulta

$$(e_1, a) = \alpha_1 (e_1, e_1) + \alpha_2' (e_1, e_2') + \ldots + \alpha_n' (e_1, e_n') = \alpha_1,$$

o sea, $(e_1, a) = 0$ cuando, y sólo cuando, $\alpha_1 = 0$.

Si el vector a pertenece al subespacio L, es decir, si $(e_1, a) = 0$, el vector $a\varphi$ también pertenecerá a L. En efecto, como la transformación φ es simétrica, resulta

$$(e_1, a\varphi) = (e_1\varphi, a) = (\lambda_0 e_1, a) = \lambda_0 (e_1, \alpha) = \lambda_0 \cdot 0 = 0,$$

o sea, el vector $a\varphi$ es ortogonal a e_1, estando por esto contenido en L. Esta propiedad del subespacio L, llamada *invariabilidad* con respecto a la transformación φ, permite considerar a φ, **aplicándola solamente a los vectores de** L, como una transformación lineal de este espacio euclídeo de $(n - 1)$ dimensiones. Esta será, incluso, una transformación simétrica del espacio L, pues, la igualdad (1), cumpliéndose para cualesquiera vectores de E_n, se cumple particularmente, para los vectores situados en L.

Por la suposición de la inducción, en el espacio L existe una base ortonormal compuesta de vectores propios de la transformación φ; designémosla mediante $e_2, \ldots, e_n$. Todos estos vectores son ortogonales al vector e_1; por consiguiente, $e_1, e_2, \ldots, e_n$ será la base ortonormal buscada del espacio E_n, que consta de vectores propios de la transformación φ. El teorema queda demostrado.

§ 37. Reducción de una forma cuadrática a los ejes principales. Par de formas

Apliquemos el último teorema del párrafo precedente para la demostración del siguiente teorema matricial:

Para cualquier matriz simétrica A se puede hallar una matriz ortogonal Q que reduzca la matriz A a la forma diagonal, o sea, que la matriz $Q^{-1}AQ$, que es la transformada de la matriz A por la matriz Q, resulta diagonal.

Sea dada una matriz simétrica A de orden n. Si $e_1, e_2, \ldots, e_n$ es una base ortonormal del espacio euclídeo E_n de n dimensiones, la matriz A determina en esta base una transformación simétrica φ. Por lo demostrado, en E_n existe una base ortonormal $f_1, f_2, \ldots, f_n$, compuesta de vectores propios de la transformación φ; en esta base, φ

se determina por una matriz diagonal B (véase el § 33). Entonces, en virtud del § 31,

$$B = Q^{-1}AQ, \tag{1}$$

donde Q es la matriz de cambio de la base f a la base e,

$$e = Qf. \tag{2}$$

Esta matriz, como matriz de cambio de una base ortonormal a otra base ortonormal, es ortogonal (véase el § 35). El teorema queda demostrado.

Como para una matriz ortogonal Q la matriz inversa es la transpuesta, $Q^{-1} = Q'$, la igualdad (1) se puede escribir en la forma

$$B = Q'AQ;$$

no obstante, por el § 26 se sabe que precisamente así se transforma la matriz simétrica A de una forma cuadrática, sometida a una transformación lineal de las indeterminadas de matriz Q. Teniendo en cuenta que una transformación lineal de las indeterminadas de matriz ortogonal es una transformación ortogonal (véase el §35) y que la forma cuadrática reducida a la forma canónica tiene una matriz diagonal, basándonos en el teorema precedente obtenemos el siguiente **teorema de reducción de una forma cuadrática real a los ejes principales:**

Toda forma cuadrática real $f(x_1, x_2, \ldots, x_n)$ se puede reducir a la forma canónica mediante una transformación ortogonal de las indeterminadas.

A pesar de que puedan existir muchas transformaciones ortogonales diferentes de las indeterminadas que reduzcan la forma cuadrática dada a la forma canónica, ésta se determina en lo fundamental unívocamente:

Cualquiera que sea la transformación ortogonal que reduzca la forma cuadrática $f(x_1, x_2, \ldots, x_n)$ de matriz A a la forma canónica, los coeficientes de esta forma canónica son las raíces características de la matriz A, tomadas con sus órdenes de multiplicidad.

En efecto, supongamos que la forma f se ha reducido ya a la forma canónica

$$f(x_1, x_2, \ldots, x_n) = \mu_1 y_1^2 + \mu_2 y_2^2 + \ldots + \mu_n y_n^2,$$

mediante una transformación ortogonal.

Esta transformación ortogonal mantiene invariable la suma de los cuadrados de las indeterminadas, de donde, si λ es una nueva indeterminada, se tiene

$$f(x_1, x_2, \ldots, x_n) - \lambda \sum_{i=1}^{n} x_i^2 = \sum_{i=1}^{n} \mu_i y_i^2 - \lambda \sum_{i=1}^{n} y_i^2.$$

Pasando a los determinantes de estas formas cuadráticas y teniendo en cuenta que después de realizar la transformación lineal el determinante de la forma cuadrática se multiplica por el cuadrado del *determinante de la transformación (véase el § 28), y que el cuadrado* del determinante de una transformación ortogonal es igual a la unidad (véase el § 35), llegamos a la igualdad

$$|A-\lambda E| = \begin{vmatrix} \mu_1-\lambda & 0 & \dots & 0 \\ 0 & \mu_2-\lambda & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & \mu_n-\lambda \end{vmatrix} = \prod_{i=1}^{n} (\mu_i-\lambda),$$

de la que se deduce la tesis del teorema.

Este resultado se puede formular también de una forma matricial:

Cualquiera que sea la matriz ortogonal que reduzca la matriz simétrica A a la forma diagonal, en la diagonal principal de la matriz diagonal obtenida figurarán las raíces características de la matriz A, tomadas con sus órdenes de multiplicidad.

Averiguación práctica de la transformación ortogonal que reduce una forma cuadrática a los ejes principales. En algunos problemas no sólo es necesario conocer la forma canónica a que se reduce una forma cuadrática real mediante una transformación ortogonal, sino también la transformación ortogonal que realiza esta reducción. Sería difícil buscar esta transformación empleando la demostración del teorema de reducción a los ejes principales, y queremos mostrar otro camino. Para esto sólo hace falta aprender a hallar la matriz ortogonal Q que reduce la matriz simétrica dada A a la forma diagonal o, lo que es lo mismo, a hallar su matriz inversa Q^{-1}. En virtud de (2), ésta es la matriz de cambio de la base e a la base f, es decir, sus filas son filas de las coordenadas (en la base e) del sistema ortonormal compuesto por n vectores propios de la transformación simétrica φ, determinada por la matriz A en la base e. No queda más que hallar tal sistema de vectores propios.

Sea λ_0 cualquier raíz característica de la matriz A y supongamos que su orden de multiplicidad es igual a k_0. Por el § 33 se sabe que el conjunto de las filas de coordenadas de todos los vectores propios de la transformación φ, correspondientes al valor propio λ_0, coincide con el conjunto de las soluciones no nulas del sistema de ecuaciones lineales homogéneas

$$(A-\lambda_0 E)\,X=0; \tag{3}$$

como la matriz A es simétrica, se puede escribir A en lugar de A'. De los teoremas de existencia de una matriz ortogonal que reduzca la matriz simétrica A a la forma diagonal, demostrados anteriormente, y de la unicidad de esta forma diagonal, se deduce que para el siste-

ma (3) se pueden hallar siempre k_0 soluciones linealmente independientes. Este sistema de soluciones se halla por los métodos conocidos en el § 12, ortogonalizando y normalizando después el sistema obtenido según el § 34.

Tomando por λ_0 cada una de las raíces características distintas de la matriz simétrica A y teniendo en cuenta que la suma de los órdenes de multiplicidad de estas raíces es igual a n, se obtiene un sistema de n vectores propios de la transformación φ, dados por sus coordenadas en la base e. Para demostrar que éste es el sistema ortonormal de los vectores propios buscados, no queda más que demostrar el siguiente **lema**:

Los vectores propios de una transformación simétrica φ que corresponden a valores propios distintos son ortogonales entre sí.

En efecto, supongamos que

$$b\varphi = \lambda_1 b, \quad c\varphi = \lambda_2 c,$$

siendo $\lambda_1 \neq \lambda_2$. Como

$$(b\varphi,\ c) = (\lambda_1 b,\ c) = \lambda_1 (b,\ c),$$
$$(b,\ c\varphi) = (b,\ \lambda_2 c) = \lambda_2 (b,\ c),$$

de

$$(b\varphi,\ c) = (b,\ c\varphi)$$

se deduce que

$$\lambda_1 (b,\ c) = \lambda_2 (b,\ c)$$

y, puesto que $\lambda_1 \neq \lambda_2$, resulta

$$(b,\ c) = 0,$$

que es lo que se quería demostrar.

Ejemplo. Reducir la forma cuadrática

$$f(x_1,\ x_2,\ x_3,\ x_4) = 2x_1x_2 + 2x_1x_3 - 2x_1x_4 - 2x_2x_3 + 2x_2x_4 + 2x_3x_4$$

a los ejes principales.

La matriz A de esta forma es

$$A = \begin{pmatrix} 0 & 1 & 1 & -1 \\ 1 & 0 & -1 & 1 \\ 1 & -1 & 0 & 1 \\ -1 & 1 & 1 & 0 \end{pmatrix}.$$

Hallemos su polinomio característico:

$$|A - \lambda E| = \begin{vmatrix} -\lambda & 1 & 1 & -1 \\ 1 & -\lambda & -1 & 1 \\ 1 & -1 & -\lambda & 1 \\ -1 & 1 & 1 & -\lambda \end{vmatrix} = (\lambda - 1)^3 (\lambda + 3).$$

Por lo tanto, la matriz A tiene la raíz característica 1 de orden tres y la raíz característica simple -3. Por consiguiente, ya se puede escribir la forma canónica a que se reduce la forma f mediante una transformación ortogonal:

$$f=y_1^2+y_2^2+y_3^2-3y_4^2.$$

Hallemos la transformación ortogonal que realiza esta reducción. El sistema de ecuaciones lineales homogéneas (3) para $\lambda_0=1$ toma la forma

$$\begin{cases} -x_1+x_2+x_3-x_4=0, \\ x_1-x_2-x_3+x_4=0, \\ x_1-x_2-x_3+x_4=0, \\ -x_1+x_2+x_3-x_4=0. \end{cases}$$

El rango de este sistema es igual a 1. Por lo tanto, para éste se pueden hallar tres soluciones linealmente independientes. Estas son, por ejemplo, los vectores

$$b_1=(1,\ 1,\ 0,\ 0),$$
$$b_2=(1,\ 0,\ 1,\ 0),$$
$$b_3=(-1,\ 0,\ 0,\ 1).$$

Ortogonalizando este sistema de vectores, se obtiene el sistema de vectores

$$c_1=b_1=(1,\ 1,\ 0,\ 0),$$
$$c_2=-\frac{1}{2}c_1+b_2=\left(\frac{1}{2},\ -\frac{1}{2},\ 1,\ 0\right),$$
$$c_3=\frac{1}{2}c_1+\frac{1}{3}c_2+b_3=\left(-\frac{1}{3},\ \frac{1}{3},\ \frac{1}{3},\ 1\right).$$

Por otra parte, el sistema de ecuaciones lineales homogéneas (3) para $\lambda_0=-3$, toma la forma

$$\begin{cases} 3x_1+x_2+x_3-x_4=0, \\ x_1+3x_2-x_3+x_4=0, \\ x_1-x_2+3x_3+x_4=0, \\ -x_1+x_2+x_3+3x_4=0. \end{cases}$$

El rango de este sistema es igual a 3. El vector

$$c_4=(1,\ -1,\ -1,\ 1).$$

es una solución no nula.

El sistema de vectores c_1, c_2, c_3, c_4 es ortogonal. Normalizando este sistema **llegamos al sistema ortonormal de vectores**

$$c_1'=\left(\frac{1}{\sqrt{2}},\ \frac{1}{\sqrt{2}},\ 0,\ 0\right),$$
$$c_2'=\left(\frac{1}{\sqrt{6}},\ -\frac{1}{\sqrt{6}},\ \sqrt{\frac{2}{3}},\ 0\right),$$
$$c_3'=\left(-\frac{1}{2\sqrt{3}},\ \frac{1}{2\sqrt{3}},\ \frac{1}{2\sqrt{3}},\ \frac{\sqrt{3}}{2}\right),$$
$$c_4'=\left(\frac{1}{2},\ -\frac{1}{2},\ -\frac{1}{2},\ \frac{1}{2}\right).$$

Por lo tanto, la forma f se reduce a los ejes principales mediante la transformación ortogonal

$$y_1=\frac{1}{\sqrt{2}}x_1+\frac{1}{\sqrt{2}}x_2,$$

$$y_2=\frac{1}{\sqrt{6}}x_1-\frac{1}{\sqrt{6}}x_2+\sqrt{\frac{2}{3}}x_3,$$

$$y_3=-\frac{1}{2\sqrt{3}}x_1+\frac{1}{2\sqrt{3}}x_2+\frac{1}{2\sqrt{3}}x_3+\frac{\sqrt{3}}{2}x_4,$$

$$y_4=\frac{1}{2}x_1-\frac{1}{2}x_2-\frac{1}{2}x_3+\frac{1}{2}x_4.$$

Es menester señalar que la elección del sistema de vectores propios, linealmente independientes, correspondientes a un valor propio múltiple, goza de mucha pluralidad; de aquí que existan muchas transformaciones ortogonales distintas que reducen la forma f a la forma conónica. Aquí sólo hemos hallado una de éstas.

Par de formas. Sea dado *un par de formas cuadráticas en n indeterminadas*, $f(x_1, x_2, \ldots, x_n)$ y $g(x_1, x_2, \ldots, x_n)$. ¿Existe alguna transformación lineal no degenerada de las indeterminadas $x_1, x_2, \ldots, x_n$ que reduzca simultáneamente ambas formas a la forma canónica?

En el caso general, la respuesta es negativa. Veamos, por ejemplo, el par de formas

$$f(x_1, x_2)=x_1^2, \quad g(x_1, x_2)=x_1x_2.$$

Supongamos que existe una transformación lineal no degenerada

$$\left.\begin{aligned} x_1&=c_{11}y_1+c_{12}y_2,\\ x_2&=c_{21}y_1+c_{22}y_2, \end{aligned}\right\} \qquad (4)$$

que reduce ambas formas a la forma canónica. Para que la forma f pueda reducirse por la transformación (4) a la forma canónica, uno de los coeficientes c_{11}, c_{12} tiene que ser igual a cero, si no aparecería el término $2c_{11}c_{12}y_1y_2$. Cambiando la numeración de las indeterminadas y_1, y_2, si esto fuese necesario, se puede suponer que $c_{12}=0$, de donde $c_{11}\neq 0$. Sin embargo, ahora resulta que

$$g(x_1, x_2)=c_{11}y_1(c_{21}y_1+c_{22}y_2)=c_{11}c_{21}y_1^2+c_{11}c_{22}y_1y_2.$$

Como también la forma g tiene que reducirse a la forma canónica, tiene que ser $c_{11}c_{22}=0$, es decir, $c_{22}=0$, lo cual, junto con $c_{12}=0$, nos lleva a lo absurdo, pues la transformación lineal (4) no es degenerada.

La situación será diferente si se supone que al menos una de nuestras formas, por ejemplo, $g(x_1, x_2, \ldots, x_n)$ es **definida positiva** *. Subsiste el teorema siguiente:

Si f y g es un par de formas cuadráticas reales en n indeterminadas, siendo la segunda de ellas definida positiva, existe una transformación lineal no degenerada que reduce simultáneamente la forma g a la forma normal y la forma f a la forma canónica.

Para la demostración, realicemos primero una transformación lineal no degenerada de las indeterminadas $x_1, x_2, \ldots, x_n$,

$$X = TY,$$

que reduzca la forma definida positiva g a la forma normal,

$$g(x_1, x_2, \ldots, x_n) = y_1^2 + y_2^2 + \ldots + y_n^2.$$

En este caso, la forma f se reducirá a otra forma φ en las nuevas indeterminadas,

$$f(x_1, x_2, \ldots, x_n) = \varphi(y_1, y_2, \ldots, y_n).$$

Realicemos ahora una transformación **ortogonal** de las indeterminadas $y_1, y_2, \ldots, y_n$,

$$Y = QZ,$$

que lleve la forma φ a los ejes principales,

$$\varphi(y_1, y_2, \ldots, y_n) = \lambda_1 z_1^2 + \lambda_2 z_2^2 + \ldots + \lambda_n z_n^2.$$

Esta transformación (véase la definición en el § 35) lleva la suma de los cuadrados de las indeterminadas $y_1, y_2, \ldots, y_n$ a la suma de los cuadrados de las indeterminadas $z_1, z_2, \ldots, z_n$. Por lo tanto, resulta

$$f(x_1, x_2, \ldots, x_n) = \lambda_1 z_1^2 + \lambda_2 z_2^2 + \ldots + \lambda_n z_n^2,$$
$$g(x_1, x_2, \ldots, x_n) = z_1^2 + z_2^2 + \ldots + z_n^2,$$

o sea, la transformación lineal

$$X = (TQ)Z$$

es la buscada.

* Claro, esta condición no es necesaria; así, pues, las formas $x_1 + x_2^2 - x_3^2$ y $x_1^2 - x_2^2 - x_3^2$ ya tienen la forma canónica, a pesar de que entre ellas no hay definidas positivas.

CAPITULO IX

CALCULO DE LAS RAICES DE LOS POLINOMIOS

§ 38. Ecuaciones de segundo, tercero y cuarto grados

El teorema fundamental demostrado en el § 23 establece la existencia de n raíces complejas para cualquier polinomio de n-ésimo grado con coeficientes numéricos. Sus demostraciones (la expuesta anteriormente, así como otras conocidas actualmente) no proporcionan, sin embargo, métodos para la averiguación práctica de estas raíces, representando «demostraciones de existencia» puras. Naturalmente, las investigaciones hechas para descubrir tales métodos comenzaron por las pruebas de deducción de fórmulas análogas a la fórmula para la resolución de la ecuación cuadrática, bien conocida por el lector para el caso de coeficientes reales en el curso escolar de álgebra. Ahora demostraremos que esta fórmula es válida también para las ecuaciones cuadráticas con coeficientes complejos, y que se pueden deducir fórmulas análogas, aunque más complicadas, para las ecuaciones de tercero y cuarto grados.

Ecuaciones cuadráticas. Sea dada la ecuación cuadrática

$$x^2+px+q=0$$

con cualesquiera coeficientes complejos; sin restringir la generalidad se puede suponer que el coeficiente superior es igual a uno. Esta ecuación se puede escribir en la forma

$$\left(x+\frac{p}{2}\right)^2+\left(q-\frac{p^2}{4}\right)=0.$$

Como es sabido, se puede extraer la raíz cuadrada del número complejo $\frac{p^2}{4}-q$ sin salir del sistema de los números complejos. Los dos valores de la raíz, que se diferencian entre sí solamente en el signo, los escribiremos en la forma $\pm\sqrt{\frac{p^2}{4}-q}$. Por lo tanto,

$$x+\frac{p}{2}=\pm\sqrt{\frac{p^2}{4}-q},$$

o sea, las raíces de la ecuación dada se pueden hallar por la fórmula ordinaria

$$x=-\frac{p}{2}\pm\sqrt{\frac{p^2}{4}-q}\,.$$

Ejempo. Resolver la ecuación

$$x^2-3x+(3-i)=0.$$

Aplicando la fórmula obtenida, resulta:

$$x=\frac{3}{2}\pm\sqrt{\frac{9}{4}-(3-i)}=\frac{3}{2}\pm\frac{1}{2}\sqrt{-3+4i}\,.$$

Por los métodos del § 19 se halla:

$$\sqrt{-3+4i}=\pm(1+2i),$$

de donde

$$x_1=2+i,\ x_2=1-i.$$

Ecuaciones cúbicas. A diferencia del caso de las ecuaciones cuadráticas, hasta ahora no tenemos un método para la resolución de las ecuaciones cúbicas, incluso en el caso de coeficientes reales. Ahora obtendremos para las ecuaciones cúbicas una fórmula análoga a la fórmula para las ecuaciones cuadráticas, suponiendo además que los coeficientes son cualesquiera números complejos.

Sea dada la ecuación cúbica

$$y^3+ay^2+by+c=0 \tag{1}$$

con coeficientes complejos cualesquiera. Sustituyendo en la ecuación (1) la incógnita y por una nueva incógnita x, ligada a y por medio de la igualdad

$$y=x-\frac{a}{3}\,, \tag{2}$$

resulta una ecuación para la incógnita x; esta ecuación, como fácilmente se comprueba, no contiene el cuadrado de esta incógnita, o sea, es una ecuación de la forma

$$x^3+px+q=0. \tag{3}$$

Hallando las raíces de la ecuación (3), en virtud de (2), se obtienen también las raíces de la ecuación (1). Por consiguiente, no queda más que aprender a resolver la ecuación cúbica «reducida» (3), con cualesquiera coeficientes complejos.

Por el teorema fundamental, la ecuación (3) posee tres raíces complejas. Sea x_0 una de estas raíces. Introduzcamos una incógnita auxiliar u y examinemos el polinomio

$$f(u)=u^2-x_0u-\frac{p}{3}\,.$$

Sus coeficientes son números complejos, poseyendo por lo tanto dos raíces complejas α y β. Por las fórmulas de Vieta:

$$\alpha+\beta=x_0, \tag{4}$$

$$\alpha\beta=-\frac{p}{3}. \tag{5}$$

Poniendo en (3) la expresión (4) de la raíz x_0, resulta

$$(\alpha+\beta)^3+p(\alpha+\beta)+q=0,$$

o bien

$$\alpha^3+\beta^3+(3\alpha\beta+p)(\alpha+\beta)+q=0.$$

Sin embargo, de (5) se deduce que $3\alpha\beta+p=0$; de donde resulta:

$$\alpha^3+\beta^3=-q. \tag{6}$$

Por otra parte, de (5) se deduce que

$$\alpha^3\beta^3=-\frac{p^3}{27}. \tag{7}$$

Las igualdades (6) y (7) muestran que los números α^3 y β^3 son raíces de la ecuación cuadrática

$$z^2+qz-\frac{p^3}{27}=0 \tag{8}$$

con coeficientes complejos.

Resolviendo la ecuación (8) se obtiene:

$$z=-\frac{q}{2}\pm\sqrt{\frac{q^2}{4}+\frac{p^3}{27}},$$

de donde*

$$\alpha=\sqrt[3]{-\frac{q}{2}+\sqrt{\frac{q^2}{4}+\frac{p^3}{27}}},\quad \beta=\sqrt[3]{-\frac{q}{2}-\sqrt{\frac{q^2}{4}+\frac{p^3}{27}}}. \tag{9}$$

Hemos obtenido la siguiente fórmula, conocida por el nombre de *fórmula de Cardano*, que expresa las raíces de la ecuación (3) mediante sus coeficientes valiéndose de radicales cuadrados y cúbicos:

$$x_0=\alpha+\beta=\sqrt[3]{-\frac{q}{2}+\sqrt{\frac{q^2}{4}+\frac{p^3}{27}}}+\sqrt[3]{-\frac{q}{2}-\sqrt{\frac{q^2}{4}+\frac{p^3}{27}}}.$$

Como el radical cúbico tiene tres valores en el campo de los números complejos, las fórmulas (9) dan tres valores para α y tres valores para β. Sin embargo, al aplicar la fórmula de Cardano, no se puede combinar cualquier valor del radical α con cualquier valor

* No importa cuál de las raíces de la ecuación (8) se toma por α^3 y cuál por β^3, puesto que en las igualdades (6) y (7), y también en la expresión (4), α y β están situadas simétricamente.

del radical β: para un valor dado de α se debe tomar solamente aquél de los tres valores de β que satisface a la condición (5).

Sea α_1 uno de los tres valores del radical α. Entonces, como se ha demostrado en el § 19, los otros dos se pueden obtener multiplicando α_1 por las raíces cúbicas ε y ε^2 de la unidad:

$$\alpha_2=\alpha_1\varepsilon,\quad \alpha_3=\alpha_1\varepsilon^2.$$

Designemos con β_1 el valor del radical β que corresponde al valor α_1 del radical α según (5), de modo que $\alpha_1\beta_1=-\frac{p}{3}$.

Los otros dos valores de β serán

$$\beta_2=\beta_1\varepsilon,\quad \beta_3=\beta_1\varepsilon^2.$$

Como $\varepsilon^3=1$,

$$\alpha_2\beta_3=\alpha_1\varepsilon\cdot\beta_1\varepsilon^2=\alpha_1\beta_1\varepsilon^3=\alpha_1\beta_1=-\frac{p}{3},$$

el valor α_2 del radical α corresponde al valor β_3 del radical β; análogamente el valor β_2 corresponde al valor α_3. Por lo tanto, todas las raíces de la ecuación (3) se pueden escribir del modo siguiente:

$$\left.\begin{aligned} x_1&=\alpha_1+\beta_1,\\ x_2&=\alpha_2+\beta_3=\alpha_1\varepsilon+\beta_1\varepsilon^2,\\ x_3&=\alpha_3+\beta_2=\alpha_1\varepsilon^2+\beta_1\varepsilon.\end{aligned}\right\}\tag{10}$$

Ecuaciones cúbicas con coeficientes reales. Veamos lo que se puede decir de las raíces de la ecuación cúbica reducida

$$x^3+px+q=0,\tag{11}$$

si sus coeficientes son reales. En este caso, desempeña un papel fundamental la expresión $\frac{q^2}{4}+\frac{p^3}{27}$ que figura en la fórmula de Cardano bajo radical cuadrado. Obsérvese que el signo de esta expresión es contrario al signo de la expresión

$$D=-4p^3-27q^2=-108\left(\frac{q^2}{4}+\frac{p^3}{27}\right),$$

denominada *discriminante* de la ecuación (11) (compárese más abajo, § 54); en las formulaciones posteriores se usará el signo del discriminante.

1) Sea $D<0$. En este caso, en la fórmula de Cardano, bajo el símbolo de cada uno de los radicales cuadrados figura un número positivo. Por esto, los números que figuran bajo los símbolos de cada uno de los radicales cúbicos son **reales**. Sin embargo, la raíz cúbica de un número real tiene un valor real y dos valores imaginarios conjugados. Sea α_1 un valor real del radical α; el valor β_1 del radical β que corresponde a α_1 por la fórmula (5), también será real, pues, el

número p también es real. Por lo tanto, resulta que la raíz $x_1 = \alpha_1 + \beta_1$ de la ecuación (11) también es real. Las otras dos raíces se hallan sustituyendo en las fórmulas (10) del presente párrafo las raíces de la unidad $\varepsilon = \varepsilon_1$ y $\varepsilon^2 = \varepsilon_2$ por sus expresiones (7) del § 19:

$$x_2 = \alpha_1\varepsilon + \beta_1\varepsilon^2 = \alpha_1\left(-\frac{1}{2} + i\frac{\sqrt{3}}{2}\right) + \beta_1\left(-\frac{1}{2} - i\frac{\sqrt{3}}{2}\right) = -\frac{\alpha_1 + \beta_1}{2} + i\sqrt{3}\,\frac{\alpha_1 - \beta_1}{2},$$

$$x_3 = \alpha_1\varepsilon^2 + \beta_1\varepsilon = \alpha_1\left(-\frac{1}{2} - i\frac{\sqrt{3}}{2}\right) + \beta_1\left(-\frac{1}{2} + i\frac{\sqrt{3}}{2}\right) = -\frac{\alpha_1 + \beta_1}{2} - i\sqrt{3}\,\frac{\alpha_1 - \beta_1}{2};$$

como los números α_1 y β_1 son reales, estas dos raíces son números imaginarios conjugados, pues, el coeficiente de la parte imaginaria es diferente de cero, debido a que $\alpha_1 \neq \beta_1$; éstos son valores de distintos radicales cúbicos.

Por lo tanto, *si $D < 0$, la ecuación (11) tiene una raíz real y dos raíces imaginarias conjugadas.*

2) Sea $D = 0$. En este caso,

$$\alpha = \sqrt[3]{-\frac{q}{2}}, \quad \beta = \sqrt[3]{-\frac{q}{2}}.$$

Sea α_1 un valor real del radical α; en virtud de (5), β_1 también será un número real, siendo $\alpha_1 = \beta_1$. Sustituyendo en las fórmulas (10), β_1 por α_1 y aplicando la igualdad $\varepsilon + \varepsilon^2 = -1$, resulta:

$$x_1 = 2\alpha_1, \quad x_2 = \alpha_1(\varepsilon + \varepsilon^2) = -\alpha_1, \quad x_3 = \alpha_1(\varepsilon^2 + \varepsilon) = -\alpha_1.$$

Por lo tanto, *si $D = 0$, todas las raíces de la ecuación (11) son reales, siendo dos de ellas iguales entre sí.*

3) Sea, finalmente, $D > 0$. En este caso, en la fórmula de Cardano, bajo el radical cuadrado figura un número real negativo y, por consiguiente, bajo los radicales cúbicos figuran números imaginarios conjugados. En consecuencia, todos los valores de los radicales α y β son ahora números imaginarios. No obstante, entre las raíces de la ecuación (11) tiene que haber por lo menos una real. Supongamos que ésta es la raíz

$$x_1 = \alpha_0 + \beta_0.$$

Como son reales tanto la suma de los números α_0 y β_0 como su producto, igual a $-\frac{p}{3}$, los números α_0 y β_0 son conjugados entre sí, pues son raíces de una ecuación cuadrática con coeficientes reales. Pero, entonces, son conjugados entre sí también los números $\alpha_0\varepsilon$ y $\beta_0\varepsilon^2$,

y también los números $\alpha_0\varepsilon^2$ y $\beta_0\varepsilon$, de donde se deduce que las raíces de la ecuación (11)

$$x_2=\alpha_0\varepsilon+\beta_0\varepsilon^2,\quad x_3=\alpha_0\varepsilon^2+\beta_0\varepsilon$$

también son números reales.

Ha resultado que las tres raíces de la ecuación (11) son reales y, además, como fácilmente se comprueba, entre ellas no hay iguales. En caso contrario, la elección de la raíz x_1 se podría realizar de modo que se cumpliese la igualdad $x_2=x_3$, de donde

$$\alpha_0(\varepsilon-\varepsilon^2)=\beta_0(\varepsilon-\varepsilon^2),$$

o sea, $\alpha_0=\beta_0$, lo cual es imposible.

Por lo tanto, *si* $D>0$, *la ecuación* (11) *tiene tres raíces reales distintas.*

El último caso que acabamos de examinar muestra que el valor práctico de la fórmula de Cardano es insignificante.

A pesar de que para $D>0$ todas las raíces de la ecuación (11) con coeficientes reales son números reales, el cálculo de éstas por la fórmula de Cardano requiere la extracción de raíces cúbicas de números imaginarios, lo que sabemos hacer solamente pasando estos números a la forma trigonométrica. Por esta razón, la expresión de las raíces mediante los radicales pierde su valor práctico. Con métodos que están fuera de los alcances de este libro, se podría demostrar que en el caso considerado las raíces de la ecuación (11) no se pueden expresar de ningún modo mediante los coeficientes, empleando radicales con expresiones reales bajo los radicales. Este caso de solución de la ecuación (11) se llama *irreducible* (¡no confundir con la irreducibilidad de los polimonios!)

Ejemplos. 1. Resolver la ecuación

$$y^3+3y^2-3y-14=0.$$

La sustitución $y=x-1$ reduce esta ecuación a la forma

$$x^3-6x-9=0. \tag{12}$$

Aquí $p=-6$, $q=-9$, por lo cual

$$\frac{q^2}{4}+\frac{p^3}{27}=\frac{49}{4}>0,$$

o sea, la ecuación (12) tiene una raíz real y dos raíces imaginarias conjugadas. Según (9),

$$\alpha=\sqrt[3]{\frac{9}{2}+\frac{7}{2}}=\sqrt[3]{8},\quad \beta=\sqrt[3]{\frac{9}{2}-\frac{7}{2}}=\sqrt[3]{1}.$$

Por consiguiente, $\alpha_1=2$, $\beta_1=1$, o sea, $x_1=3$. Las otras dos raíces se hallan por las fórmulas (10):

$$x_2=-\frac{3}{2}+i\,\frac{\sqrt{3}}{2},\quad x_3=-\frac{3}{2}-i\,\frac{\sqrt{3}}{2}.$$

De aquí se deduce que las raíces de la ecuación dada son:

$$y_1=2,\ y_2=-\frac{5}{2}+i\frac{\sqrt{3}}{2},\ y_3=-\frac{5}{2}-i\frac{\sqrt{3}}{2}.$$

2. Resolver la ecuación

$$x^3-12x+16=0.$$

Aquí $p=-12$, $q=16$, por lo tanto,

$$\frac{q^2}{4}+\frac{p^3}{27}=0.$$

De aquí se deduce que $\alpha=\sqrt[3]{-8}$, o sea, $\alpha_1=-2$. En consecuencia,

$$x_1=-4,\ x_2=x_3=2.$$

3. Resolver la ecuación

$$x^3-19x+30=0.$$

Aquí $p=-19$, $q=30$, de donde

$$\frac{q^2}{4}+\frac{p^3}{27}=-\frac{784}{27}<0.$$

Así, manteniéndose en el campo de los números reales, la fórmula de Cardano no es aplicable a pesar de que sus raíces son los números reales 2, 3 y -5.

Ecuaciones de cuarto grado. La resolución de la ecuación de cuarto grado

$$y^4+ay^3+by^2+cy+d=0 \tag{13}$$

con coeficientes complejos arbitrarios se reduce a la resolución de una ecuación cúbica auxiliar. Esto se consigue con el método siguiente, perteneciente a Ferrari.

Se reduce previamente la ecuación (13) con la sustitución $y=x-\frac{a}{4}$, a la forma

$$x^4+px^2+qx+r=0. \tag{14}$$

Luego, el primer miembro de esta ecuación se transforma idénticamente mediante el parámetro auxiliar α, del modo siguiente:

$$x^4+px^2+qx+r=\left(x^2+\frac{p}{2}+\alpha\right)^2+qx+r-\frac{p^2}{4}-\alpha^2-2\alpha x^2-p\alpha,$$

o bien

$$\left(x^2+\frac{p}{2}+\alpha\right)^2-\left[2\alpha x^2-qx+\left(\alpha^2+p\alpha-r+\frac{p^2}{4}\right)\right]=0. \tag{15}$$

Elijamos ahora α de modo que el polinomio que figura entre corchetes sea un cuadrado completo. Para esto, es necesario que tenga una raíz múltiple, es decir, se tiene que cumplir la igualdad

$$q^2-4\cdot 2\alpha\left(\alpha^2+p\alpha-r+\frac{p^2}{4}\right)=0. \tag{16}$$

La igualdad (16) es una ecuación cúbica con respecto a la incógnita α con coeficientes complejos. Como se sabe, esta ecuación tiene tres raíces complejas. Sea α_0 una de ellas; en virtud de la fórmula de Cardano, ésta se *expresa por radicales mediante los coeficientes de la* ecuación (16), o sea, mediante los coeficientes de la ecuación (14).

Con tal elección del valor de α, el polinomio que figura entre *corchetes en (15) tiene una raíz múltiple* $\frac{q}{4\alpha_0}$, *de segundo orden. Por* consiguiente, la ecuación (15) toma la forma

$$\left(x^2+\frac{p}{2}+\alpha_0\right)^2-2\alpha_0\left(x-\frac{q}{4\alpha_0}\right)^2=0,$$

es decir, se descompone en dos ecuaciones cuadráticas:

$$\left.\begin{aligned} x^2-\sqrt{2\alpha_0}x+\left(\frac{p}{2}+\alpha_0+\frac{q}{2\sqrt{2\alpha_0}}\right)=0,\\ x^2+\sqrt{2\alpha_0}x+\left(\frac{p}{2}+\alpha_0-\frac{q}{2\sqrt{2\alpha_0}}\right)=0. \end{aligned}\right\} \qquad (17)$$

Como las ecuaciones (17) se han obtenido de la ecuación (14) haciendo transformaciones idénticas, las raíces de las ecuaciones (17) serán también raíces de la ecuación (14). Fácilmente se observa también que las raíces de la ecuación (14) se expresan por radicales mediante los coeficientes. Aquí no escribiremos las fórmulas correspondientes, pues son muy complicadas y prácticamente inútiles; tampoco estudiaremos separadamente el caso en que la ecuación (14) tenga coeficientes reales.

Observación sobre las ecuaciones de grado superior. A pesar de que los griegos ya conocían los métodos de resolución de las ecuaciones cuadráticas, el descubrimiento de los métodos de resolución de las ecuaciones de tercero y cuarto grado, expuesto anteriormente, pertenece al siglo XVI. Durante casi tres siglos se continuaron haciendo estériles pruebas para dar el paso siguiente, es decir, para hallar fórmulas que expresasen las raíces de cualquier ecuación de quinto grado (o sea, de una ecuación de quinto grado con coeficientes *literales*) *mediante sus coeficientes por radicales. Estas pruebas terminaron* en los años veinte del siglo pasado, después de que Abel demostró que no existen tales fórmulas para las ecuaciones de n-ésimo grado, cuando $n \geqslant 5$.

Sin embargo, el resultado de Abel no excluía la posibilidad de que las raíces de cualquier polinomio concreto con coeficientes numéricos se pudiesen expresar de algún modo mediante los coeficientes empleando *alguna combinación de radicales o, como está convenido decir,* que cualquier ecuación se resolviese por radicales. El problema sobre las condiciones según las cuales una ecuación dada es resoluble por radicales fue estudiado detalladamente por Galois en los años

treinta del siglo pasado. Resultó que para cualquier n, empezando desde $n = 5$, se pueden indicar ecuaciones de n-ésimo grado irresolubles por radicales, que tienen incluso coeficientes numéricos enteros. Tal es, por ejemplo, la ecuación

$$x^5 - 4x - 2 = 0.$$

Las investigaciones de Galois influyeron definitivamente en el desarrollo del álgebra. Sin embargo, nuestra tarea no incluye su exposición.

§ 39. Acotación de las raíces

Ya sabemos que no existe un método para calcular los valores exactos de las raíces de los polinomios con coeficientes numéricos. No obstante, diversos problemas de la mecánica, de la física y de la técnica se reducen al problema de las raíces de los polinomios, los cuales suelen ser a veces de grados suficientemente altos. Esta circunstancia fue la causa de numerosas investigaciones que tenían por objeto aprender a hacer tales o cuales deducciones sobre las raíces de los polinomios con coeficientes numéricos sin conocer estas raíces. Se estudiaba, por ejemplo, la cuestión sobre la posición de las raíces en el plano complejo (las condiciones según las cuales todas las raíces están dentro del círculo unidad, o sea, que en su valor absoluto son menores que la unidad, o bien, las condiciones para que todas las raíces estén situadas en el semiplano izquierdo, o sea, que sus partes reales fuesen negativas, etc). Para los polinomios de coeficientes reales, se elaboraban métodos de definición del número de raíces reales, se buscaban las cotas entre las que podían estar estas raíces, etc. Finalmente, fueron dedicadas muchas investigaciones a los métodos del cálculo aproximado de las raíces. Ordinariamente, en las aplicaciones técnicas es suficiente conocer solamente los valores aproximados de las raíces con cierta exactitud prefijada, y si, por ejemplo, las raíces del polinomio se expresasen incluso por radicales, éstos se sustituirían de todos modos por sus valores aproximados.

En su tiempo, todas estas investigaciones formaban el contenido fundamental del álgebra superior. En nuestro curso está incluida solamente una parte muy pequeña de los resultados relacionados con esto y, teniendo en cuenta las necesidades primarias de las aplicaciones, nos limitaremos al caso de polinomios de coeficientes reales y de sus raíces reales, saliéndonos pocas veces de estos límites. Además, se va a considerar sistemáticamente el polinomio $f(x)$ de coeficientes reales como una función real (continua) de la variable real x. Siempre que sea útil se emplearán los resultados y métodos del análisis matemático.

Es conveniente comenzar el estudio de las raíces reales de un polinomio $f(x)$ de coeficientes reales considerando la gráfica de este polinomio. Evidentemente, *las raíces reales del polinomio son las abscisas de los puntos de intersección de su gráfica con el eje x, y sólo éstas.*

Veamos, por ejemplo, el polinomio de quinto grado

$$h(x) = x^5 + 2x^4 - 5x^3 + 8x^2 - 7x - 3.$$

Por los resultados del § 24, sobre las raíces de este polinomio, se puede afirmar lo siguiente: como es de grado impar, $h(x)$ tiene por lo menos una raíz real; si el número de raíces reales es mayor que uno, será igual a tres o a cinco, pues las raíces imaginarias son conjugadas a pares.

El estudio de la gráfica del polinomio $h(x)$ permite afirmar algo más sobre sus raíces. Tracemos esta gráfica (fig. 9) *, considerando sólo los valores enteros de x y calculando los valores correspondientes de $h(x)$ por el método de Horner:

x	$h(x)$
.	.
.	.
-4	-39
-3	144
-2	83
-1	18
0	-3
1	-4
2	39
.	.
.	.

Fig. 9.

Vemos, pues, que el polinomio $h(x)$ posee al menos tres raíces reales: una positiva α_1 y dos negativas α_2 y α_3, siendo

$$1 < \alpha_1 < 2, \quad -1 < \alpha_2 < 0,$$
$$-4 < \alpha_3 < -3.$$

La información sobre las raíces (reales) del polinomio, obtenida al examinar la gráfica, suele ser prácticamente bastante buena. Sin embargo, siempre quedan algunas dudas: no se sabe si verdaderamente se han hallado todas las raíces o no. Así, en el ejemplo consi-

* En el dibujo, en el eje y se ha tomado una escala diez veces menor que en el eje x.

derado no se ha demostrado que a la derecha del punto $x = 2$ y a la izquierda del punto $x = -4$ ya no hay raíces del polinomio. Además, como sólo se han tomado valores enteros de x, se puede suponer que la gráfica trazada refleja con poca exactitud el comportamiento de la función $h(x)$, pueda ser incluso que no tenga en cuenta algunas de sus más pequeñas oscilaciones, perdiéndose así algunas raíces.

Claro, al construir la gráfica se podrían tomar no sólo los valores enteros de x, sino también valores que se diferenciasen en 0,1 o en 0,01. Sin embargo, con esto se complicarían considerablemente los cálculos de los valores de $h(x)$, y de todos modos persistirían las dudas indicadas anteriormente. Por otra parte, con los métodos del análisis matemático se podrían hallar los máximos y mínimos de la función $h(x)$ y comparar nuestra gráfica con el comportamiento verdadero de la función; pero esto trae consigo la cuestión sobre las raíces de la derivada $h'(x)$, o sea, el mismo problema que estamos resolviendo.

De aquí surge la necesidad de métodos más perfectos para la búsqueda de cotas entre las que están comprendidas las raíces reales de un polinomio de coeficientes reales, y la determinación del número de estas raíces. Ahora nos vamos a ocupar del problema sobre las **cotas de las raíces reales,** dejando para los siguientes párrafos la cuestión sobre la cantidad de estas raíces.

La demostración del lema sobre el módulo del término superior (véase el § 23) proporciona ya una cota para los módulos de las raíces de un polinomio. En efecto, haciendo $k = 1$ en la desigualdad (3) del § 23, resulta que, para

$$|x| \geqslant 1 + \frac{A}{|a_0|}, \tag{1}$$

donde a_0 es el coeficiente superior y A, el máximo de los módulos de los demás coeficientes, el módulo del término superior del polinomio es mayor que el módulo de la suma de todos los demás términos. Por consiguiente, ningún valor de x que satisfaga a la desigualdad (1) puede ser raíz de este polinomio.

Por lo tanto, *para un polinomio $f(x)$ con cualesquiera coeficientes numéricos, el número $1 + \frac{A}{a_0}$ es una cota superior para los módulos de todas sus raíces, reales o imaginarias.* Así, pues, para el polinomio $h(x)$ examinado más arriba, esta cota es el número 9, puesto, que $a_0 = 1$, $A = 8$.

No obstante, esta cota suele ser demasiado grande, si sólo nos interesan las cotas de las raíces reales. Ahora se expondrán otros métodos más exactos. Hay que tener presente que a pesar de que se marquen las cotas entre las que tienen que estar comprendidas las raíces reales del polinomio, **esto no significa que tales raíces existan.**

Demostremos primero que *es suficiente conocer la cota superior de las raíces positivas de cualquier polinomio.* En efecto, sea dado un polinomio $f(x)$ de grado n y sea N_0 una cota superior de sus raíces positivas. Examinemos los polinomios

$$\varphi_1(x) = x^n f\left(\frac{1}{x}\right),$$

$$\varphi_2(x) = f(-x),$$

$$\varphi_3(x) = x^n f\left(-\frac{1}{x}\right)$$

y hallemos las cotas superiores de sus raíces positivas; supongamos que éstas son los números N_1, N_2, y N_3, respectivamente. Entonces, *el número* $\frac{1}{N_1}$ *será una cota inferior de las raíces positivas del polinomio* $f(x)$, pues, si α es una raíz positiva de $f(x)$, $\frac{1}{\alpha}$ es una raíz positiva de $\varphi_1(x)$, y de $\frac{1}{\alpha} < N_1$, resulta $\alpha > \frac{1}{N_1}$. Análogamente, *los números* $-N_2$ *y* $-\frac{1}{N_3}$ *son las cotas inferior y superior, respectivamente, de las raíces negativas del polinomio* $f(x)$. Por lo tanto, todas las raíces positivas del polinomio $f(x)$ satisfacen a las desigualdades $\frac{1}{N_1} < x < N_0$ y todas las raíces negativas, satisfacen a las desigualdades

$$-N_2 < x < -\frac{1}{N_3}.$$

Para determinar una cota superior de las raíces positivas se puede aplicar el método siguiente. Sea dado un polinomio

$$f(x) = a_0x^n + a_1x^{n-1} + \ldots + a_n$$

con coeficientes reales, siendo $a_0 > 0$. Supongamos ahora que a_k, $k \geqslant 1$, es el primer coeficiente negativo; si no hubiese tales coeficientes, el polinomio $f(x)$ no podría tener raíces positivas. Sea, finalmente, B el máximo valor absoluto de los coeficientes negativos. Entonces *el número*

$$1 + \sqrt[k]{\frac{B}{a_0}}$$

es una cota superior de las raíces positivas del polinomio $f(x)$.

En efecto, suponiendo $x > 1$ y sustituyendo cada uno de los coeficientes $a_1, a_2, \ldots, a_{k-1}$ por cero y cada uno de los coeficientes $a_k, a_{k+1}, \ldots, a_n$ por B, se puede disminuir solamente el valor

del polinomio, resultando

$$f(x) \geqslant a_0x^n - B(x^{n-k} + x^{n-k-1} + \ldots + x + 1) = a_0x^n - B\,\frac{x^{n-k+1}-1}{x-1}$$

y, como $x > 1$,

$$f(x) > a_0x^n - \frac{Bx^{n-k+1}}{x-1} = \frac{x^{n-k+1}}{x-1}\,[a_0x^{k-1}(x-1) - B]. \qquad (2)$$

Si

$$x > 1 + \sqrt[k]{\frac{B}{a_0}}, \qquad (3)$$

entonces, como

$$a_0x^{k-1}(x-1) - B \geqslant a_0(x-1)^k - B,$$

la expresión que figura entre corchetes en la fórmula (2) resulta positiva, sea, en virtud de (2), el valor de $f(x)$ es estrictamente positivo. Por lo tanto, los valores de x que satisfacen a la desigualdad (3) no pueden ser raíces de $f(x)$, como se quería demostrar.

Para el polinomio $h(x)$ considerado anteriormente, como $k = 2$ y $B = 7$, para la cota superior de las raíces positivas este método da el número $1 + \sqrt{7}$, que se puede sustituir por el número entero próximo mayor 4.

De los numerosos métodos existentes de acotación superior de las raíces positivas expondremos solamente el *método de Newton*. Este método es más complicado que el expuesto anteriormente, pero, no obstante, da ordinariamente, muy buen resultado.

Sea dado un polinomio $f(x)$ de coeficientes reales y con el coeficiente superior positivo a_0. *Si para $x = c$, el polinomio $f(x)$ y todas sus derivadas sucesivas $f'(x)$, $f''(x)$, . . ., $f^{(n)}(x)$ toman valores positivos, el número c es una cota superior de las raíces positivas.*

En efecto, según la fórmula de Taylor (véase el § 23),

$$f(x) = f(c) + (x-c)f'(c) + (x-c)^2\frac{f''(c)}{2!} + \ldots + (x-c)^n\frac{f^{(n)}(c)}{n!}.$$

Vemos, que si $x \geqslant c$, el segundo miembro será un número estrictamente positivo, es decir, tales valores de x no pueden ser raíces de $f(x)$.

Al buscar el número correspondiente c, para un polinomio $f(x)$ dado, es conveniente obrar del modo siguiente. La derivada $f^{(n)}(x) = n!a_0$ es un número positivo, de donde, el polinomio $f^{(n-1)}(x)$ es una función creciente de x. Por consiguiente, existe un número c_1 tal, que para $x \geqslant c_1$ la derivada $f^{(n-1)}(x)$ es positiva. De esto se deduce que para $x \geqslant c_1$, la derivada $f^{(n-2)}(x)$ es una función creciente de x, por lo cual, existe un número c_2 tal $(c_2 \geqslant c_1)$, que para $x \geqslant c_2$, la derivada $f^{(n-2)}(x)$ también es positiva. Continuando de este modo, llegaremos por fin al número buscado c.

Apliquemos el método de Newton al polinomio $h(x)$ examinado anteriormente.

$$h(x)=x^5+2x^4-5x^3+8x^2-7x-3,$$
$$h'(x)=5x^4+8x^3-15x^2+16x-7,$$
$$h''(x)=20x^3+24x^2-30x+16,$$
$$h'''(x)=60x^2+48x-30,$$
$$h^{IV}(x)=120x+48,$$
$$h^{V}(x)=120.$$

Fácilmente se comprueba (aunque sea por el método de Horner), que todos estos polinomios son positivos para $x=2$. Por lo tanto, el número 2 es una cota superior de las raíces positivas del polinomio $h(x)$, resultado que es mucho más exacto que los obtenidos por otros métodos.

Para hallar una cota inferior de las raíces negativas del polinomio $h(x)$, veamos el polinomio $\varphi_2(x)=-h(-x)$*. Como

$$\varphi_2(x)=x^5-2x^4-5x^3-8x^2-7x+3,$$
$$\varphi_2'(x)=5x^4-8x^3-15x^2-16x-7,$$
$$\varphi_2''(x)=20x^3-24x^2-30x-16,$$
$$\varphi_2'''(x)=60x^2-48x-30,$$
$$\varphi_2^{IV}(x)=120x-48,$$
$$\varphi^{V}(x)=120,$$

y todos estos polinomios son positivos para $x=4$, lo que fácilmente se comprueba, el número 4 es una cota superior de las raíces positivas de $\varphi_2(x)$, de donde, el número -4 es una cota inferior de las raíces negativas de $h(x)$.

Examinando, finalmente, los polinomios

$$\varphi_1(x)=-x^5h\left(\frac{1}{x}\right)=3x^5+7x^4-8x^3+5x^2-2x-1,$$
$$\varphi_3(x)=-x^5h\left(-\frac{1}{x}\right)=3x^5-7x^4-8x^3-5x^2-2x+1,$$

y aplicando de nuevo el método de Newton, para las cotas superiores de las raíces positivas de estos polinomios hallamos los números 1 y 4, respectivamente; de aquí el número $\frac{1}{1}=1$ es una cota inferior de las raíces positivas del polinomio $h(x)$; el número $-\frac{1}{4}$ es una cota superior de las raíces negativas de éste.

Por lo tanto, las raíces positivas del polinomio $h(x)$ están comprendidas entre los números 1 y 2, y las raíces negativas, entre los números -4 y $-\frac{1}{4}$. Este resultado concuerda perfectamente con lo hallado antes al examinar la *gráfica*.

---

* Aquí tomamos $-h(-x)$ en lugar de $h(-x)$, porque para la aplicación del método de Newton, el coeficiente superior tiene que ser positivo. Naturalmente, este cambio de signo no influye en las raíces del polinomio $\varphi_2(x)$.

§ 40. Teorema de Sturm

Ahora estudiaremos el problema sobre **el número de raíces reales que tiene un polinomio $f(x)$ de coeficientes reales.** Mas, nos interesará tanto el número total de las raíces reales como los números de las raíces positivas y negativas por separado y, en general, el número de raíces comprendidas entre dos números dados a y b. Existen unos cuantos métodos para la averiguación del número exacto de raíces, siendo éstos demasiado complicados; entre ellos, el más sensillo es el *método de Sturm* que se expondrá a continuación.

Introduzcamos primero una definición que se utilizará también en el párrafo siguiente.

Sea dado un sistema finito ordenado de números reales diferentes de cero, por ejemplo,

$$1,\ 3,\ -2,\ 1,\ -4,\ -8,\ -3,\ 4,\ 1. \tag{1}$$

Escribamos sucesivamente los signos de estos números:

$$+,\ +,\ -,\ +,\ -,\ -,\ -,\ +,\ +. \tag{2}$$

Observamos que en el sistema (2) figuran cuatro veces signos contrarios consecutivos. En virtud de esto, se dice que el sistema ordenado (1) presenta cuatro *variaciones de signo*. Naturalmente, **el número de variaciones de signo** puede ser calculado para cualquier sistema finito ordenado de números reales diferentes de cero.

Consideremos ahora un polinomio $f(x)$ de coeficientes reales y supongamos que éste **carece de raíces múltiples,** pues, en caso contrario, se le podría dividir por el máximo común divisor del mismo y su derivada. Un sistema finito ordenado de polinomios, no nulos, de coeficientes reales

$$f(x)=f_0(x),\ f_1(x),\ f_2(x),\ \ldots,\ f_s(x) \tag{3}$$

se llama *sistema de Sturm* del polinomio $f(x)$ si se cumplen las condiciones siguientes:

1) Los polinomios consecutivos del sistema (3) no tienen raíces comunes.

2) El último polinomio $f_s(x)$ no tiene raíces reales.

3) Si α es una raíz real de uno de los polinomios intermedios $f_k(x)$ del sistema (3), $1\leqslant k\leqslant s-1$, entonces, $f_{k-1}(\alpha)$ y $f_{k+1}(\alpha)$ tienen diferente signo.

4) Si α es una raíz real del polinomio $f(x)$, el producto $f(x)\,f_1(x)$ cambia su signo de menos a más, cuando al crecer x pasa por el punto α.

El problema de la existencia de un sistema de Sturm para cualquier polinomio se estudiará más adelante; ahora, suponiendo que $f(x)$

posea tal sistema, señalaremos el modo de utilizarlo para averiguar el número de raíces reales.

Si el número real c no es raíz del polinomio dado $f(x)$ y (3) es el sistema de Sturm de este polinomio, tomamos el sistema de números reales

$$f(c),\ f_1(c),\ f_2(c)\ \ldots,\ f_s(c),$$

eliminamos en éste todos los números iguales a cero, y designamos con $W(c)$ el número de variaciones de signo que presenta el sistema obtenido; diremos que $W(c)$ *es el número de variaciones de signo que presenta el sistema de Sturm* (3) *del polinomio* $f(x)$ *para* $x = c$ *.

Subsiste el siguiente

Teorema de Sturm. *Si los números reales* a *y* b, $a < b$, *no son raíces del polinomio* $f(x)$, *el cual carece de raíces múltiples, entonces* $W(a) \geqslant W(b)$, *y la diferencia* $W(a) - W(b)$ *es igual al número de raíces reales del polinomio* $f(x)$ *comprendidas entre* a *y* b.

Por lo tanto, para la determinación del número de raíces reales del polinomio $f(x)$ comprendidas entre a y b (recordamos que, por hipótesis, $f(x)$ no tiene raíces múltiples), sólo hay que averiguar en cuánto disminuye el número de variaciones de signo que presenta el sistema de Sturm de este polinomio al pasar del valor a al valor b.

Para la demostración del teorema, veamos cómo cambia el número $W(x)$ al crecer x. Mientras x no pase por alguna raíz de alguno de los polinomios del sistema de Sturm (3), los signos de los polinomios de este sistema no cambiarán y no variará el número $W(x)$. En virtud de esto, y también debido a la condición 2) de la definición del sistema de Sturm, no queda más que examinar dos casos: el paso de x por una raíz de uno de los polinomios intermedios $f_k(x)$, $1 \leqslant k \leqslant s-1$, y el paso de x por una raíz del mismo polinomio $f(x)$.

Sea α una raíz del polinomio $f_k(x)$, $1 \leqslant k \leqslant s-1$. Entonces, por la condición 1), $f_{k-1}(\alpha)$ y $f_{k+1}(\alpha)$ son diferentes de cero. Por consiguiente, se podrá hallar un número positivo ε, posiblemente muy pequeño, de tal modo que en el intervalo $(\alpha - \varepsilon, \alpha + \varepsilon)$ los polinomios $f_{k-1}(x)$ y $f_{k+1}(x)$ no tengan raíces, conservando por ello constantes los signos, que serán además **distintos,** por la condición (3). De esto se deduce que cada uno de los sistemas de números

$$f_{k-1}(\alpha-\varepsilon),\ f_k(\alpha-\varepsilon),\ f_{k+1}(\alpha-\varepsilon), \tag{4}$$

$$f_{k-1}(\alpha+\varepsilon),\ f_k(\alpha+\varepsilon),\ f_{k+1}(\alpha+\varepsilon) \tag{5}$$

prensentan exactamente una variación de signo, independientemente de los signos que tengan los números $f_k(\alpha - \varepsilon)$ y $f_k(\alpha + \varepsilon)$. Por

* Naturalmente, las variaciones de signo que presenta el sistema de Sturm de un polinomio $f(x)$ no tiene nada de común con la variación de signo del mismo polinomio $f(x)$, debida al paso de x por una raíz de este polinomio.

ejemplo, si en el intervalo considerado $f_{k-1}(x)$ es negativo y $f_{k+1}(x)$ es positivo, y si $f_k(\alpha - \varepsilon) > 0$, $f_k(\alpha + \varepsilon) < 0$, a los sistemas (4) y (5) les corresponderán los sistemas de signos:

$$-, +, +; -, -, +.$$

Por lo tanto, al pasar x por una raíz de uno de los polinomios intermedios del sistema de Sturm, la variación de signos en este sistema sólo puede trasladarse, mas no podrá aparecer de nuevo ni desaparecer, por lo que **durante tal paso el número $W(x)$ no variará.**

Supongamos, por otra parte, que α es una raíz del mismo polinomio $f(x)$. Según la condición 1), en este caso α no será raíz de $f_1(x)$. Por consiguiente, existe un número positivo ε tal, que el intervalo $(\alpha - \varepsilon, \alpha + \varepsilon)$ no contiene raíces del polinomio $f_1(x)$, por lo cual este último mantiene constante el signo en este intervalo. Si este signo es positivo, en virtud de la condición 4), al pasar x por α, el mismo polinomio $f(x)$ cambia el signo de menos a más, es decir, $f(\alpha - \varepsilon) < 0$, $f(\alpha + \varepsilon) > 0$. Luego, a los sistemas de números

$$f(\alpha-\varepsilon),\ f_1(\alpha-\varepsilon) \ \text{y}\ f(\alpha+\varepsilon),\ f_1(\alpha+\varepsilon) \tag{6}$$

les corresponden los sistemas de signos

$$-, + \ \text{y}\ +, +,$$

o sea, en el sistema de Sturm **se pierde una variación.** Si el signo de $f_1(x)$ es negativo en el intervalo $(\alpha - \varepsilon, \alpha + \varepsilon)$, de nuevo en virtud de la condición 4), el polinomio $f(x)$ cambia el signo de más a menos al pasar x por α, o sea, $f(\alpha - \varepsilon) > 0$, $f(\alpha + \varepsilon) < 0$; a los sistemas de números (6) les corresponden ahora los sistemas de signos

$$+, - \ \text{y}\ -, -,$$

es decir, en el sistema de Sturm **se pierde de nuevo una variación.**

Por lo tanto, *el número $W(x)$ varía (al crecer x) solamente cuando x pasa por una raíz del polinomio $f(x)$, disminuyendo exactamente, en este caso, en una unidad.*

Naturalmente, con esto queda demostrado el teorema de Sturm. Para aplicar este teorema a la averiguación del número total de raíces reales de un polinomio $f(x)$, es suficiente tomar por a el límite inferior de las raíces negativas y por b, el límite superior de las raíces positivas. Sin embargo, es más fácil obrar del modo siguiente. En virtud del lema demostrado en el § 23, existe un número positivo N, posiblemente muy grande, tal que para $|x| > N$ los signos **de todos** los polinomios del sistema de Sturm coinciden con los signos de sus términos superiores. En otras palabras, existe un valor positivo tan grande de la indeterminada x, que los signos de los valores correspondientes de todos los polinomios del sistema de Sturm coinciden con los signos de sus **coeficientes** superiores; este valor

de x, cuyo cálculo no es necesario, se designa convencionalmente con el símbolo ∞. Por otra parte, existe un número negativo x, cuyo valor absoluto es tan grande que los signos de los valores correspondientes de los polinomios del sistema de Sturm coinciden con los signos de sus coeficientes superiores para los polinomios de grado par, y son contrarios a los signos de los coeficientes superiores para los polinomios de grado impar; convengamos en designar este valor de x mediante $-\infty$. Está claro que en el intervalo $(-\infty, \infty)$ están contenidas todas las raíces reales de todos los polinomios del sistema de Sturm y, en particular, todas las raíces reales del polinomio $f(x)$. Aplicando el teorema de Sturm a este intervalo, se halla el número de estas raíces; la aplicación del teorema de Sturm a los intervalos $(-\infty, 0)$ y $(0, \infty)$ proporciona el número de raíces negativas y el número de raíces positivas del polinomio $f(x)$, respectivamente.

No queda más que demostrar que *cualquier polinomio $f(x)$ de coeficientes reales que no tenga raíces múltiples posee un sistema de Sturm*. Entre los diversos métodos que se emplean para la construcción de tal sistema expondremos el más usual. Hagamos $f_1(x) = f'(x)$, con lo que se garantiza el cumplimiento de la condición 4) de la definición del sistema de Sturm. En efecto, si α es una raíz real del polinomio $f(x)$, se tiene $f'(\alpha) \neq 0$. Si $f'(\alpha) > 0$, entonces $f'(x) > 0$ en un entorno del punto α. Por lo tanto, al pasar x por α, $f(x)$ cambia el signo de menos a más; esto mismo se cumple también para el producto $f(x) f_1(x)$. Razonamientos análogos son válidos también para el caso en que $f'(\alpha) < 0$. Se divide luego $f(x)$ por $f_1(x)$ y el residuo de esta división, **tomado con signo contrario,** se toma por $f_2(x)$:

$$f(x) = f_1(x) q_1(x) - f_2(x).$$

En general, si ya se han hallado los polinomios $f_{k-1}(x)$ y $f_k(x)$, el polinomio $f_{k+1}(x)$ será el residuo de la división de $f_{k-1}(x)$ por $f_k(x)$, tomado con signo contrario:

$$f_{k-1}(x) = f_k(x) q_k(x) - f_{k+1}(x). \tag{7}$$

El método expuesto se diferencia del algoritmo de Euclides, aplicado a los polinomios $f(x)$ y $f'(x)$, solamente en que cada vez se cambia el signo al residuo, y la división consiguiente se efectúa ya por este residuo, tomado con el signo contrario. Como al buscar el máximo común divisor este cambio de signos no importa, nuestro proceso terminará en cierto $f_s(x)$, que será el máximo común divisor de los polinomios $f(x)$ y $f'(x)$; además, como $f(x)$ no tiene raíces múltiples, o sea, es primo con $f'(x)$, resulta que en realidad $f_s(x)$ será un número real diferente de cero.

De aquí que el sistema construido de polinomios

$$f(x) = f_0(x), \ f'(x) = f_1(x), \ f_2(x), \ \ldots, \ f_s(x)$$

también satisface a la condición 2) de la definición del sistema de Sturm. Para demostrar que se cumple la condición 1), supongamos que los polinomios consecutivos $f_k(x)$ y $f_{k+1}(x)$ tienen una raíz común α. Entonces, por la igualdad (7), α también es raíz del polinomio $f_{k-1}(x)$. Pasando a la igualdad

$$f_{k-2}(x)=f_{k-1}(x)\,q_{k-1}(x)-f_k(x),$$

resulta que α es también raíz de $f_{k-2}(x)$. Continuando de este modo, hallamos que α es una raíz común de $f(x)$ y $f'(x)$, lo que contradice a las hipótesis hechas. Finalmente, el cumplimiento de la condición 3) es consecuencia inmediata de la igualdad (7), pues, si $f_k(\alpha)=0$, resulta $f_{k-1}(\alpha)=-f_{k+1}(\alpha)$.

Apliquemos el método de Sturm al polinomio

$$h(x)=x^5+2x^4-5x^3+8x^2-7x-3.$$

examinado en el párrafo anterior. Aquí no comprobaremos previamente que $h(x)$ carece de raíces múltiples, puesto que el método de construcción del sistema de Sturm, sirve a la vez para comprobar si el polinomio y su derivada son primos entre sí.

Hallemos el sistema de Sturm para $h(x)$ aplicando el método indicado. Mas, a diferencia del algoritmo de Euclides, en el proceso de división multiplicaremos y simplificaremos solamente por números **positivos** arbitrarios, puesto que los signos de los residuos desempeñan un papel fundamental en el método de Sturm. Obtendremos el sistema siguiente:

$$h(x)=x^5+2x^4-5x^3+8x^2-7x-3.$$
$$h_1(x)=5x^4+8x^3-15x^2+16x-7,$$
$$h_2(x)=66x^3-150x^2+172x+61,$$
$$h_3(x)=-464x^2+1135x+723,$$
$$h_4(x)=-32\,599\,457x-8\,486\,093,$$
$$h_5(x)=-1.$$

Determinemos los signos de los polinomios de este sistema para $x=-\infty$ y $x=\infty$, para lo cual, según lo indicado, se deben observar solamente los signos de los coeficientes superiores y los grados de estos polinomios. Resulta la tabla:

	$h(x)$	$h_1(x)$	$h_2(x)$	$h_3(x)$	$h_4(x)$	$h_5(x)$	Número de variaciones de signo
$-\infty$	$-$	$+$	$-$	$-$	$+$	$-$	4
∞	$+$	$+$	$+$	$-$	$-$	$-$	1

Por lo tanto, al pasar x de $-\infty$ a ∞, el sistema de Sturm pierde tres variaciones de signo y, por esto, el polinomio $h(x)$ tiene exactamente tres raíces

reales. De aquí vemos que, al construir la gráfica de este polinomio en el párrafo anterior, no habíamos perdido ninguna raíz.

Apliquemos el método de Sturm a otro polinomio más simple. Sea dado el polinomio

$$f(x) = x^3 + 3x^2 - 1.$$

Hallemos el número de sus raíces reales y también las cotas enteras entre las que está comprendida cada una de estas raíces, sin construir previamente la gráfica del mismo.

El sistema de Sturm de este polinomio es

$$f(x) = x^3 + 3x^2 - 1,$$
$$f_1(x) = 3x^2 + 6x,$$
$$f_2(x) = 2x + 1,$$
$$f_3(x) = 1.$$

Hallemos el número de variaciones de signo que presenta este sistema para $x = -\infty$ y $x = \infty$.

	$f(x)$	$f_1(x)$	$f_2(x)$	$f_3(x)$	Número de variaciones de signo
$-\infty$	$-$	$+$	$-$	$+$	3
∞	$+$	$+$	$+$	$+$	0

Por consiguiente, el polinomio $f(x)$ tiene tres raíces reales. Para determinar más exactamente la posición de estas raíces, continuemos la tabla anterior:

	$f(x)$	$f_1(x)$	$f_2(x)$	$f_3(x)$	Número de variacones de signo
$x=-3$	$-$	$+$	$-$	$+$	3
$x=-2$	$+$	0	$-$	$+$	2
$x=-1$	$+$	$-$	$-$	$+$	2
$x=0$	$-$	0	$+$	$+$	1
$x=1$	$+$	$+$	$+$	$+$	0

Por lo tanto, el sistema de Sturm del polinomio $f(x)$ pierde una variación de signo cada vez que x pasa de *-3 a -2*, de *-1 a 0* y de *0 a 1*. *Luego*, las raíces α_1, α_2 y α_3 del polinomio satisfacen a las desigualdades:

$$-3 < \alpha_1 < -2, \quad -1 < \alpha_2 < 0, \quad 0 < \alpha_3 < 1.$$

§ 41. Otros teoremas sobre el número de raíces reales

El teorema de Sturm resuelve por completo el problema del número de raíces reales de un polinomio. No obstante, su defecto fundamental consiste en que los cálculos necesarios para la construcción del sistema de Sturm son muy engorrosos, de lo cual se puede convencer el lector realizando todos los cálculos respectivos en el primero de los ejemplos considerados anteriormente. En virtud de esto, se demostrarán ahora dos teoremas que no proporcionarán el número exacto de raíces reales, sino solamente una cota **superior** de este número. Después de haber hallado mediante la gráfica una cota **inferior** para el número de raíces reales, la aplicación de estos teoremas dará la posibilidad de hallar a veces el número exacto de raíces reales sin recurrir al método de Sturm.

Sea dado un polinomio $f(x)$ de n-ésimo grado, de coeficientes reales, que puede tener raíces múltiples. Consideremos el sistema formado por sus derivadas sucesivas

$$f(x)=f^{(0)}(x),\ f'(x),\ f''(x),\ \ldots,\ f^{(n-1)}(x),\ f^{(n)}(x), \tag{1}$$

en el cual la última es igual al coeficiente superior a_0 del polinomio $f(x)$, multiplicado por $n!$, conservando por consiguiente constante el signo. Si el número real c no es raíz de ninguno de los polinomios del sistema (1), el número de variaciones de signo que presenta el sistema ordenado de números

$$f(c),\ f'(c),\ f''(c),\ \ldots,\ f^{(n-1)}(c),\ f^{(n)}(c).$$

se designará con $S(c)$.

Por lo tanto, se puede considerar la función $S(x)$, definida para los valores de x que no anulan a ninguno de los polinomios del sistema (1).

Veamos cómo varía el número $S(x)$ al crecer x. Mientras x no pase por una raíz de alguno de los polinomios (1), el número $S(x)$ no puede variar. En virtud de esto, tenemos que examinar dos casos: el paso de x por una raíz del polinomio $f(x)$ y el paso de x por una raíz de una de las derivadas $f^{(k)}(x)$, $1\leqslant k\leqslant n-1$.

Sea α una raíz múltiple de orden l del polinomio $f(x)$, $l\geqslant 1$, o sea,

$$f(\alpha)=f'(\alpha)=\ldots=f^{(l-1)}(\alpha)=0,\ f^{(l)}(\alpha)\neq 0.$$

Sea ε un número positivo tan pequeño que el intervalo $(\alpha-\varepsilon,\ \alpha+\varepsilon)$ no contenga raíces de los polinomios $f(x)$, $f'(x)$, ..., $f^{(l-1)}(x)$ diferentes de α y no contenga tampoco ninguna raíz del polinomio $f^{(l)}(x)$. Demostremos que en el sistema de números

$$f(\alpha-\varepsilon),\ f'(\alpha-\varepsilon),\ \ldots,\ f^{(l-1)}(\alpha-\varepsilon),\ f^{(l)}(\alpha-\varepsilon),$$

dos números consecutivos cualesquiera tienen signos contrarios, mientras que todos los números

$$f(\alpha+\varepsilon),\ f'(\alpha+\varepsilon),\ \ldots,\ f^{(l-1)}(\alpha+\varepsilon),\ f^{(l)}(\alpha+\varepsilon)$$

tienen un mismo signo. Como cada uno de los polinomios del sistema (1) es la derivada del polinomio anterior, no nos queda más que demostrar que si x pasa por una raíz α del polinomio $f(x)$, entonces, independientemente del orden de multiplicidad de esta raíz, $f(x)$ y $f'(x)$ tenían signos contrarios antes del paso, y después del paso sus signos coinciden. Si $f(\alpha-\varepsilon)>0$, entonces $f(x)$ decrece en el intervalo $(\alpha-\varepsilon,\ \alpha)$, de donde, $f'(\alpha-\varepsilon)<0$; si $f(\alpha-\varepsilon)<0$, entonces $f(x)$ crece y, por lo tanto, $f'(\alpha-\varepsilon)>0$. Por consiguiente, en ambos casos los signos son distintos. Por otra parte, si $f(\alpha+\varepsilon)>0$, entonces $f(x)$ crece en el intervalo $(\alpha,\ \alpha+\varepsilon)$ y $f'(\alpha+\varepsilon)>0$; análogamente, de $f(\alpha+\varepsilon)<0$ se deduce que $f'(\alpha+\varepsilon)<0$. Por lo tanto, después de pasar por la raíz α, los signos de $f(x)$ y $f'(x)$ tienen que coincidir.

De lo demostrado se deduce que al pasar x por una raíz de orden l del polinomio $f(x)$, el sistema

$$f(x),\ f'(x),\ \ldots,\ f^{(l-1)}(x),\ f^{(l)}(x)$$

pierde l variaciones de signo.

Sea ahora α una raíz de las derivadas

$$f^{(k)}(x),\ f^{(k+1)}(x),\ \ldots,\ f^{(k+l-1)}(x),\ 1\leqslant k\leqslant n-1,\ l\geqslant 1,$$

no siendo raíz de $f^{(k-1)}(x)$ y tampoco de $f^{(k+l)}(x)$. Por lo demostrado anteriormente, el paso de x por α da lugar a una pérdida de l variaciones de signo en el sistema:

$$f^{(k)}(x),\ f^{(k+1)}(x),\ \ldots,\ f^{(k+l-1)}(x),\ f^{(k+l)}(x).$$

Por cierto, este paso puede crear una nueva variación de signo entre $f^{(k-1)}(x)$ y $f^{(k)}(x)$; sin embargo, como $l\geqslant 1$, al pasar x por α, el número de variaciones de signo en el sistema

$$f^{(k-1)}(x),\ f^{(k)}(x),\ f^{(k+1)}(x),\ \ldots,\ f^{(k+l-1)}(x),\ f^{(k+l)}(x),$$

o no varía, o disminuye. Pero, puede disminuir solamente en un número par, pues los polinomios $f^{(k-1)}(x)$ y $f^{(k+l)}(x)$ no cambian sus signos al pasar x por el valor α.

De los resultados obtenidos se deduce que, *si los números a y b, $a<b$, no son raíces de ninguno de los polinomios del sistema* (1), *el número de raíces reales del polinomio $f(x)$, comprendidas entre a y b y contadas cada una de ellas tantas veces como lo indique su orden de multiplicidad, es igual a la diferencia $S(a)-S(b)$ o es menor que esta diferencia en un número par.*

Para debilitar las restricciones impuestas a los números a y b, introduzcamos las siguientes notaciones. Supongamos que el número real c no es raíz del polinomio $f(x)$, pudiendo ser, posiblemente, raíz de otros polinomios del sistema (1). Designemos con $S_+(c)$ el número de variaciones de signo que presenta el sistema de números

$$f(c),\ f'(c),\ f''(c),\ \ldots,\ f^{(n-1)}(c),\ f^{(n)}(c) \tag{2}$$

calculado del modo siguiente: si

$$f^{(k)}(c) = f^{(k+1)}(c) = \ldots = f^{(k+l-1)}(c) = 0, \tag{3}$$

pero

$$f^{(k-1)}(c) \neq 0,\ f^{(k+l)}(c) \neq 0, \tag{4}$$

entonces se supone que $f^{(k)}(c)$, $f^{(k+1)}(c)$, . . ., $f^{(k+l+1)}(c)$ tienen el mismo signo que $f^{(k+1)}(c)$; por supuesto, esto equivale a suponer que al calcular el número de variaciones de signo que presenta el sistema (2), los ceros se han eliminado. Por otra parte, designemos con $S_-(c)$ el número de variaciones de signo que presenta el sistema (2), calculado del modo siguiente: cumpliéndose las condiciones (3) y (4), se supone que $f^{(k+i)}(c)$, $0 \leqslant i \leqslant l-1$, tiene el mismo signo que $f^{(k+l)}(c)$, si la diferencia $l-i$ es par, y el signo contrario, si esta diferencia es impar.

Si se quiere determinar ahora el número de raíces reales del polinomio $f(x)$, comprendidas entre a y b, $a < b$, donde a y b no son raíces de $f(x)$, pudiendo ser, posiblemente, raíces de otros polinomios del sistema (1), se obra del modo siguiente. Sea ε tan pequeño que el intervalo $(a, a+2\varepsilon)$ no contenga raíces del polinomio $f(x)$ y tampoco raíces diferentes de a de los demás polinomios del sistema (1); sea, por otra parte, η tan pequeño que el intervalo $(b-2\eta, b)$ no contenga raíces de $f(x)$ y tampoco raíces, diferentes de b, de los demás polinomios del sistema (1). Entonces, el número buscado de raíces reales del polinomio $f(x)$ será igual al número de raíces reales de este polinomio, comprendidas entre $a+\varepsilon$ y $b-\eta$, o sea, por lo demostrado anteriormente, será igual a la diferencia $S(a+\varepsilon) - S(b-\eta)$ o será menor que esta diferencia en un número par. Mas, fácilmente se observa que

$$S(a+\varepsilon) = S_+(a),\ S(b-\eta) = S_-(b).$$

Con esto queda demostrado el siguiente

Teorema de Budan — Fourier. *Si los números reales a y b, $a < b$, no son raíces del polinomio $f(x)$ de coeficientes reales, el número de raíces reales de este polinomio, comprendidas entre a y b, y contadas cada una de ellas tantas veces como indique su orden de multiplicidad, es igual a la diferencia $S_+(a) - S_-(b)$ o es menor que esta diferencia en un número par.*

Designemos con el símbolo ∞ un valor positivo tan grande de la indeterminada x que los signos de los valores correspondientes de todos los polinomios del sistema (1) coincidan con los signos de sus coeficientes superiores. Como estos coeficientes son sucesivamente los números a_0, na_0, $n(n-1)a_0$, ..., $n!a_0$, cuyos signos coinciden, resulta $S(\infty) = S_-(\infty) = 0$. Por otra parte, como

$$f(0) = a_n,\ f'(0) = a_{n-1},\ f''(0) = a_{n-2}2!,$$
$$f'''(0) = a_{n-3}3!,\ \ldots,\ f^{(n)}(0) = a_0 \cdot n!,$$

donde $a_0, a_1, \ldots, a_n$ son los coeficientes del polinomio $f(x)$, resulta que $S_+(0)$ coincide con el número de variaciones de signo que presenta el sistema de coeficientes del polinomio $f(x)$, en el cual no se cuentan los coeficientes iguales a cero. Así, aplicando el teorema de Budan—Fourier al intervalo $(0, \infty)$, resulta el teorema siguiente:

Teorema de Descartes. *El número de raíces positivas de un polinomio $f(x)$, contadas cada una tantas veces como indique su orden de multiplicidad, es igual al número de variaciones de signo que presenta el sistema de coeficientes de este polinomio (los coeficientes iguales a cero no se cuentan) o es menor que este número en un número par.*

Está claro que para la determinación del número de raíces negativas del polinomio $f(x)$ es suficiente aplicar el teorema de Descartes al polinomio $f(-x)$. Naturalmente, si en este caso ninguno de los coeficientes del polinomio $f(x)$ es igual a cero, a las variaciones de signo que presenta el sistema de coeficientes del polinomio $f(-x)$ corresponden **permanencias** de signo que presenta el sistema de coeficientes del polinomio $f(x)$, y viceversa. Por lo tanto, *si un polinomio $f(x)$ no tiene coeficientes iguales a cero, el número de sus raíces negativas (contadas con su orden de multiplicidad) es igual al número de permanencias de signo que presenta el sistema de coeficientes o es menor que éste en un número par.*

He aquí *otra demostración más del teorema de Descartes* que no se basa en el teorema de Budan — Fourier. Demostremos primero el lema siguiente:

Si $c > 0$, entonces el número de variaciones de signo que presenta el sistema de coeficientes del polinomio $f(x)$, es menor en un número impar que el número de variaciones de signo que presenta el sistema de los coeficientes del producto $(x-c)f(x)$.

En efecto, encerrando entre paréntesis los términos consecutivos de un mismo signo, expresemos el polinomio $f(x)$, cuyo coeficiente superior a_0 se supone positivo, del modo siguiente:

$$f(x) = (a_0x^n + \ldots + b_1x^{k_1+1}) - (a_1x_1^{k_1} + \ldots + b_2x^{k_2+1}) + \ldots$$
$$\ldots + (-1)^s(a_sx^{k_s} + \ldots + b_{s+1}x^l). \qquad (5)$$

Aquí $a_0 > 0$, $a_1 > 0$, . . ., $a_s > 0$, mientras que b_1, b_2, . . ., b_s son positivos o iguales a cero; pero b_{s+1} se supone estrictamente positivo, de modo que x^t, donde $t \geqslant 0$, es la potencia mínima de la indeterminada x que figura en el polinomio $f(x)$, con un coeficiente diferente de cero. La expresión

$$a_0x^n + \ldots + b_1x^{k_1+1})$$

puede constar eventualmente de un solo sumando; esto sucede cuando $k_1 + 1 = n$. Observaciones análogas se refieren también a otras expresiones entre paréntesis que figuran en la fórmula (5).

Escribamos ahora el polinomio igual al producto $(x - c) f(x)$, en el que separaremos solamente los términos que contengan las potencias $n + 1$, $k_1 + 1$, . . ., $k_s + 1$ y t de la indeterminada x. Resulta

$$(x-c)\, f(x) = (a_0x^{n+1} + \ldots) - (a_1'x^{k_1+1} + \ldots) + \ldots$$
$$\ldots + (-1)^s (a_s'x^{k_s+1} + \ldots - cb_{s+1}x^t), \quad (6)$$

donde $a_i' = a_i + cb_i$, $i = 1, 2, \ldots, s$, por lo cual, como $c > 0$, todas las a_i' son estrictamente positivas. Por lo tanto, el sistema de coeficientes del polinomio $f(x)$ presenta entre los términos a_0x^n y $-a_1x^{k_1}$ (y también entre los términos $-a_1x^{k_1}$ y $a_2x^{k_2}$, etc). una variación de signo, y el polinomio $(x - c)\, f(x)$ presenta entre los términos correspondientes a a_0x^{n+1} y $-a_1'x^{k_1+1}$ (respectivamente, entre los términos $-a_1'x^{k_1+1}$ y $a_2x^{k_2+1}$, etc.) una variación de signo o más, pero en este último caso, inevitablemente, en **un número par** más. Aquí no nos interesan los lugares exactos de estas variaciones de signo; por ejemplo, puede ocurrir que el coeficiente de x^{k_1+2} en (6) sea negativo, igual que el coeficiente $-a_1'$ y que, por esto, estos dos coeficientes consecutivos no presenten variación de signo, es decir, que entre los paréntesis primeros las variaciones de signo estén situadas antes. Obsérvese ahora que los últimos paréntesis en (5) no presentaban ninguna variación de signo, mientras que los últimos paréntesis en (6) si presentan, y, además, un número **impar** de ellas. Téngase en cuenta que los últimos coeficientes de los polinomios $f(x)$ y $(x - c)\, f(x)$, diferentes de cero, o sea, $(-1)^s\, b_{s+1}$ y $(-1)^{s+1}b_{s+1}c$, tienen signos contrarios. Por lo tanto, al pasar de $f(x)$ a $(x - c)\, f(x)$ el número total de variaciones de signo que presenta el sistema de coeficientes aumenta inevitablemente en un número impar (naturalmente, la suma de unos cuantos términos, de los cuales uno es impar y los demás son pares, es impar). El lema está demostrado.

Para demostrar el teorema de Descartes, designemos con $\alpha_1, \alpha_2, \ldots, \alpha_k$ todas las raíces positivas del polinomio $f(x)$.

Por lo tanto,

$$f(x) = (x-\alpha_1)(x-\alpha_2)\ldots(x-\alpha_h)\varphi(x),$$

donde $\varphi(x)$ es un polinomio de coeficientes reales sin raíces reales positivas. De aquí se deduce que el primero y el último coeficiente del polinomio $\varphi(x)$, diferentes de cero, son de un mismo signo, o sea, el sistema de coeficientes de este polinomio presenta un número par de variaciones de signo. Aplicando ahora sucesivamente el lema demostrado anteriormente a los polinomios

$$\varphi(x),\ (x-\alpha_1)\varphi(x),\ (x-\alpha_1)(x-\alpha_2)\varphi(x),\ \ldots,\ f(x),$$

se obtiene que el número de variaciones de signo que presenta el sistema de coeficientes aumenta cada vez en un número impar, o sea, en una unidad más un número par; por esto, el número de variaciones de signo que presenta el sistema de coeficientes del polinomio $f(x)$ es mayor que el número k en un número par.

Apliquemos los teoremas de Descartes y de Budan-Fourier al polinomio *examinado anteriormente*:

$$h(x) = x^5 + 2x^4 - 5x^3 + 8x^2 - 7x - 3.$$

El *número de variaciones de signo* que presenta el sistema de coeficientes es igual a tres y, por consiguiente, según el teorema de Descartes, $h(x)$ puede tener una o tres raíces positivas. Por otra parte, $h(x)$ no tiene coeficientes iguales a cero, y como el sistema de coeficientes presenta dos permanencias de signo, resulta que $h(x)$ o bien tiene dos raíces negativas, o bien, no tiene ninguna. Comparando con los resultados obtenidos antes mediante la gráfica, vemos que dos es el número exacto de raíces negativas de nuestro polinomio.

Para la determinación exacta del número de raíces positivas, aplicaremos el teorema de Budan-Fourier al intervalo $(1, \infty)$, pues, en el § 39 ya se había demostrado que 1 es una cota inferior de las raíces positivas del polinomio $h(x)$. Las derivadas sucesivas de $h(x)$ también fueron halladas en el § 39. Hallemos sus signos para $x = 1$ y $x = \infty$:

	$h(x)$	$h'(x)$	$h''(x)$	$h'''(x)$	$h^{IV}(x)$	$h^{V}(x)$	Número de variaciones de signo
$x=1$	−	+	+	+	+	+	1
$x=\infty$	+	+	+	+	+	+	0

De aquí se deduce, que el sistema de derivadas, al pasar x de 1 a ∞, pierde una variación de signo, por lo que, $h(x)$ tiene exactamente una raíz positiva.

Obsérvese que, en general, al buscar el número de raíces reales de un polinomio, se debe comenzar con la construcción de la gráfica y aplicar los teoremas de Descartes y Budan-Fourier; solamente en casos muy extremos se debe pasar a construir el sistema de Sturm.

El teorema de Descartes se puede precisar cuando se sabe previamente que todas las raíces del polinomio son reales, como esto tiene lugar, por ejemplo, en el caso del polinomio característico de una matriz simétrica. Resulta que:

Si todas las raíces del polinomio $f(x)$ son reales y el término independiente es diferente de cero, el número k_1 de raíces positivas de este polinomio es igual al número s_1 de variaciones de signo que presenta el sistema de sus coeficientes, y el número k_2 de raíces negativas es igual al número s_2 de variaciones de signo que presenta el sistema de coeficientes del polinomio $f(-x)$.

En efecto, en estas condiciones,

$$k_1 + k_2 = n, \tag{7}$$

donde n es el grado del polinomio $f(x)$, y, según el teorema de Descartes

$$k_1 \leqslant s_1, \quad k_2 \leqslant s_2. \tag{8}$$

Demostremos que

$$s_1 + s_2 \leqslant n. \tag{9}$$

La demostración se hará por el método de inducción sobre n, puesto que, como $a_0 \neq 0$, $a_1 \neq 0$, para $n = 1$ presenta variación de signo solamente uno de los polinomios

$$f(x) = a_0x + a_1, \quad f(-x) = -a_0x + a_1,$$

o sea, en este caso $s_1 + s_2 = 1$. Supongamos que la fórmula (9) ya está demostrada para los polinomios de grado menor que n. Si

$$f(x) = a_0x^n + a_{n-l}x^l + \ldots + a_n,$$

donde $l \leqslant n-1$, $a_{n-l} \neq 0$, hacemos

$$g(x) = a_{n-l}x^l + \ldots + a_n.$$

Entonces

$$f(x) = a_0x^n + g(x), \quad f(-x) = (-1)^n a_0x^n + g(-x).$$

Si s_1' y s_2' son los números de variaciones de signo que presentan los sistemas de coeficientes de los polinomios $g(x)$ y $g(-x)$, respectivamente, entonces, según la hipótesis de inducción (claro que $l \geqslant 1$),

$$s_1' + s_2' \leqslant l.$$

Si $l = n-1$, entonces, solamente uno de los polinomios $f(x)$ o $f(-x)$ presentará una variación de signo en el primer sitio, o sea, para $f(x)$, entre a_0 y $a_1 = a_{n-l}$; por consiguiente

$$s_1 + s_2 = s_1' + s_2' + 1 \leqslant l + 1 = n.$$

Si $l \leqslant n-2$, entonces, cada uno de los polinomios $f(x)$, $f(-x)$ puede presentar variaciones de signo en los primeros lugares, pero, en este caso,

$$s_1 + s_2 \leqslant s_1' + s_2' + 2 \leqslant l + 2 \leqslant (n-2) + 2 = n.$$

Confrontando (7), (8) y (9), se obtiene que

$$k_1 = s_1, \quad k_2 = s_2,$$

como se quería demostrar.

§ 42. Cálculo aproximado de las raíces

Los métodos expuestos en los párrafos anteriores permiten efectuar la *separación* de las raíces reales de un polinomio $f(x)$ de coeficientes reales, es decir, indicar para cada raíz las cotas entre las que la raíz está comprendida. Si estas cotas son bastante estrechas, cualquier número comprendido entre ellas se puede tomar por valor aproximado de la raíz buscada. Por lo tanto, después de establecer, por el método de Sturm (o por otro método más sencillo), que entre los números **racionales** a y b está comprendida una sola raíz del polinomio $f(x)$, se plantea el problema de aproximar estas cotas entre sí, de modo que las nuevas cotas a' y b' tengan un número prefijado de sus primeras cifras decimales iguales; con esto, la raíz buscada quedará calculada con la exactitud dada.

Existen muchos métodos que permiten hallar con suficiente rapidez el valor aproximado de la raíz con la exactitud deseada. Aquí se indicarán dos de ellos, los que teóricamente son más simples y generales; al aplicarlos simultáneamente se obtiene el resultado con una rapidez satisfactoria. Es menester observar que los métodos que se van a exponer, no sólo pueden aplicarse a los polinomios, sino también a clases más amplias de funciones continuas.

A continuación se supondrá que α es una raíz **simple** del polinomio $f(x)$ (ya que podemos librarnos siempre de las raíces múltiples) y que la raíz α ya está separada de las demás raíces por las cotas a y b, $a < \alpha < b$; en particular, de aquí se deduce que $f(a)$ y $f(b)$ tienen signo contrario.

Método de interpolación lineal (llamado también *regula falsi*). Por valor aproximado de la raíz α se podría tomar, por ejemplo, la semisuma de las cotas a y b, $\frac{a+b}{2}$, o sea, el punto medio del intervalo limitado por los puntos a y b. Sin embargo, es más natural suponer que la raíz está más cerca de la cota que corresponde a un valor absoluto menor del polinomio. El método de interpolación lineal consiste en que se toma por valor aproximado de la raíz α el número c que divide el intervalo (a, b) en partes proporcionales

a los valores absolutos de los números $f(a)$ y $f(b)$, o sea,

$$\frac{c-a}{b-c} = -\frac{f(a)}{f(b)};$$

el signo menos del segundo miembro es debido a que $f(a)$ y $f(b)$ tienen signos contrarios. De aquí que

$$c = \frac{bf(a) - af(b)}{f(a) - f(b)}. \tag{1}$$

Como muestra la fig. 10, el método de interpolación lineal consiste en que en el intervalo (a, b) la curva $y = f(x)$ se sustituye por la cuerda que une los puntos $A(a, f(a))$ y $B(b, f(b))$, tomando por valor aproximado de la raíz α la abscisa del punto de intersección de esta cuerda con el eje x.

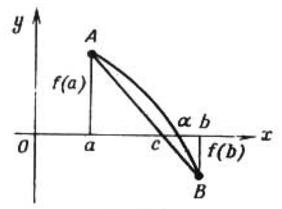


Fig. 10.

Método de Newton. Como α es una raíz simple del polinomio $f(x)$, se tiene $f'(\alpha) \neq 0$. Supongamos que también $f''(\alpha) \neq 0$, pues, en caso contrario, el problema se reduciría al cálculo de la raíz del polinomio $f''(x)$, que es de menor grado que $f(x)$. Supongamos que el intervalo (a, b) no contiene raíces de $f(x)$ diferentes de α, ni contiene tampoco ninguna raíz del polinomio $f'(x)$ y del

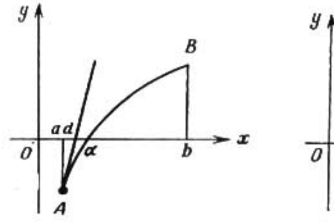


Fig. 11.

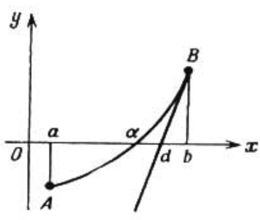


Fig. 12.

polinomio $f''(x)$ *. Por lo tanto, como se deduce del curso de análisis matemático, en el intervalo (a, b) la curva $y = f(x)$ es monótona creciente, o es monótona decreciente; además, en todos los puntos

* El estrechamiento de las cotas que da lugar a que se satisfaga esta condición se consigue ordinariamente sin dificultad alguna, pues los métodos expuestos anteriormente permiten determinar el número de raíces de los polinomios $f'(x)$ y $f''(x)$ en cualquier intervalo.

de este intervalo la convexidad está dirigida hacia arriba, o en todos los puntos la convexidad está dirigida hacia abajo. Por consiguiente, en la representación de la curva en el intervalo (a, b) pueden presentarse cuatro casos, expuestos en las figs. 11—14.

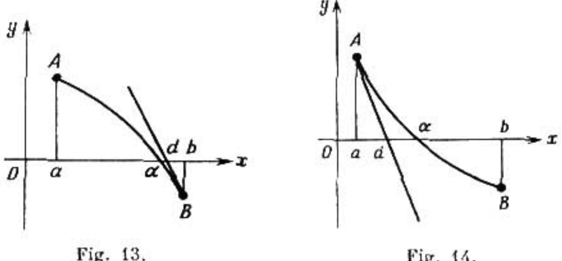


Fig. 13. Fig. 14.

Designemos con a_0 uno de los extremos a o b, en el que **el signo de $f(x)$ coincide con el signo de $f''(x)$**. Como $f(a)$ y $f(b)$ tienen signos distintos y $f''(x)$ conserva el signo en todo el intervalo (a, b), tal a_0 puede ser indicado. En los casos representados en las figs. 11 y 14,

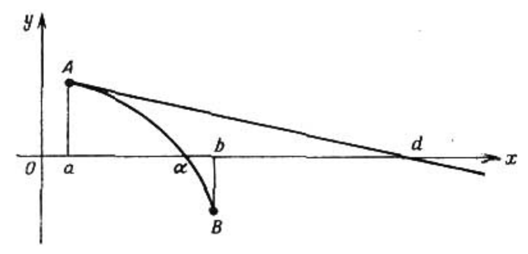


Fig. 15.

$a_0 = a$; en los otros dos casos, $a_0 = b$. Tracemos por el punto de abscisa a_0, es decir, por el punto de coordenadas $(a_0, f(a_0))$, la tangente a la curva $y = f(x)$ y designemos con d la abscisa del punto de intersección de esta tangente con el eje x. Las figs. 11—14 muestran que el número d se puede tomar por valor aproximado de la raíz α. Por consiguiente, el método de Newton consiste en sustituir la curva $y = f(x)$ en el intervalo (a, b) por su tangente, trazada en uno de los extremos de este intervalo. La condición impuesta a la elección del punto a_0 es esencial: la fig. 15 muestra que omitiendo

esta condición el punto de intersección de la tangente con el eje x puede estar muy lejos de ser una aproximación de la raíz buscada.

Hallemos la fórmula según la cual se busca el número d. Como se sabe, la ecuación de la tangente a la curva $y = f(x)$ en el punto $(a_0, f(a_0))$ se puede escribir en la forma

$$y - f(a_0) = f'(a_0)(x - a_0).$$

Poniendo aquí las coordenadas $(d, 0)$ del punto de intersección de la tangente con el eje x, resulta

$$-f(a_0) = f'(a_0)(d - a_0),$$

de donde

$$d = a_0 - \frac{f(a_0)}{f'(a_0)}. \qquad (2)$$

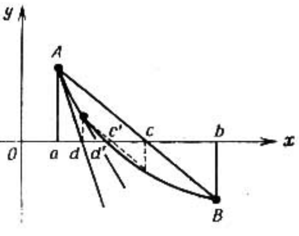


Fig. 16.

Si el lector une en las figs. 11—14 los puntos A y B con cuerdas, observará que *en todos los casos los métodos de interpolación lineal y de Newton dan una aproximación al valor verdadero de la raíz* α *por lados diversos.* Por esto, si el intervalo (a, b) satisface a las condiciones que se piden en el método de Newton, es conveniente **combinar** estos dos métodos. De este modo se obtienen para la raíz unas cotas más estrechas: c y d. Si éstas no dan todavía la exactitud de aproximación pedida, se les pueden aplicar otra vez más a estos límites ambos métodos (véase la fig. 16), etc; además, se puede demostrar que este proceso permite calcular verdaderamente la raíz α con la exactitud que se desee.

Apliquemos estos métodos al polinomio

$$h(x) = x^5 - 2x^4 - 5x^3 + 8x^2 - 7x - 3,$$

considerado en los párrafos anteriores.

Ya sabemos que este polinomio tiene una raíz simple α_1 comprendida entre los límites $1 < \alpha_1 < 2$. Se puede afirmar previamente, que estas cotas son demasiado amplias para que los métodos de interpolación lineal y de Newton, aplicados una sola vez, puedan dar un buen resultado. Sin embargo, los aplicaremos para tener un ejemplo de cálculos poco complicados.

Como ya vimos en el párrafo anterior, para $x = 1$ las derivadas $h'(x)$, $h''(x)$, ..., $h^{V}(x)$ toman valores positivos. Basándose en los resultados del § 39, se deduce que el valor $x = 1$ es, para $h'(x)$, y también para $h''(x)$, una cota superior de las raíces positivas. Por consiguiente, el intervalo $(1, 2)$ no contiene raíces de estas derivadas, pudiéndose aplicarle el método de Newton. Además, $h''(x)$ es positiva en todo este intervalo, y como

$$h(1) = -4, \quad h(2) = 39,$$

hay que poner $a_0 = 2$. Teniendo en cuenta que $h'(2) = 109$, aplicando la fórmula (2), hallamos:

$$d = 2 - \frac{39}{109} = \frac{179}{109} = 1{,}64\ldots$$

Por otra parte, la fórmula (1) da

$$c = \frac{2\cdot(-4)-1\cdot 39}{-4-39} = \frac{47}{43} = 1{,}09\ldots$$

y, por consiguiente, la raíz α_1 está comprendida entre las cotas

$$1{,}09 < \alpha_1 < 1{,}65.$$

Hemos obtenido un estrechamiento de las cotas demasiado insignificante para que este resultado sea satisfactorio. Claro, a las nuevas cotas obtenidas se les podrían aplicar de nuevo nuestros métodos. Sin embargo, sería conveniente hallar desde el principio para α_1 unas cotas bastante estrechas, por ejemplo, con una exactitud de 0,1 e incluso hasta de 0,01, y solamente después aplicar estos métodos. Naturalmente, esto daría lugar a que los cálculos se complicasen muchísimo, pero al resolver problemas concretos, en los que se necesitan conocer las raíces de un polinomio con bastante exactitud, no hay más remedio que actuar de este modo.

Volvamos a examinar nuestro polinomio $h(x)$ y su raíz α_1. Obsérvese que todos los valores de los polinomios que aparecen a continuación se calculan por la regla de Horner. Como

$$h(1{,}3) = -0{,}13987,\quad h(1{,}31) = 0{,}0662923851,$$

se tiene

$$1{,}3 < \alpha_1 < 1{,}31,$$

es decir, hemos hallado el valor de la raíz α_1 con una exactitud de 0,01. Apliquemos ahora los métodos de interpolación lineal a estas nuevas cotas:

$$c = \frac{1{,}31\cdot(-0{,}13987)-1{,}3\cdot 0{,}0662923851}{-0{,}13987-0{,}0662923851} = \frac{0{,}26940980063}{0{,}2061623851} = 1{,}30678\ldots$$

Apliquemos el método de Newton a estas mismas cotas, en donde se debe poner $a_0 = 1{,}31$. Como

$$h'(1{,}31) = 20{,}92822405,$$

se tiene

$$d = 1{,}31 - \frac{0{,}0662923851}{20{,}92822405} = \frac{27{,}3496811204}{20{,}92822405} = 1{,}30683\ldots$$

Por lo tanto

$$1{,}30678 < \alpha_1 < 1{,}30684,$$

por consiguiente, poniendo $\alpha_1 = 1{,}30681$, se comete un error menor que 0,00003.

Hasta ahora no hemos demostrado que los métodos expuestos anteriormente permiten calcular la raíz con la exactitud deseada, o sea, no hemos demostrado la convergencia de estos métodos. Demostremos esto únicamente para el método de Newton.

Supongamos de nuevo que la raíz simple α del polinomio $f(x)$ está contenida en el intervalo (a, b), siendo éste elegido de la forma necesaria para la aplicación del método de Newton. De aquí se deduce, en particular, la existencia de unos números positivos A y B tales, que en todo el intervalo (a, b)

$$|f'(x)| > A,\quad |f''(x)| < B. \tag{3}$$

Hagamos la notación

$$C=\frac{B}{2A}$$

y supongamos que

$$C(b-a)<1. \qquad (4)$$

Para que se cumpla esta desigualdad, habrá posiblemente que sustituir las cotas (a, b) de la raíz α por otras más estrechas, lo cual no influye para que se cumplan las desigualdades (3). Sea a_0 la cota a o b, a la que se debe aplicar el método de Newton. Aplicando la fórmula (2), por valores aproximados de la raíz α obtenemos, sucesivamente, los números $a_1, a_2, \ldots, a_k, \ldots$, situados en el intervalo (a, b) y relacionados entre sí por las igualdades

$$a_k=a_{k-1}-\frac{f(a_{k-1})}{f'(a_{k-1})},\ k=1,\ 2,\ \ldots \qquad (5)$$

Sea

$$\alpha=a_k+h_k,\ k=0,\ 1,\ 2,\ \ldots \qquad (6)$$

Entonces

$$0=f(\alpha)=f(a_k)+h_kf'(a_k)+\frac{h_k^2}{2}f''(a_k+\theta h_k),$$

donde $0<\theta<1$. Debido a las condiciones impuestas al intervalo (a, b), $f'(a_k)\neq 0$, y teniendo en cuenta (5) y (6), resulta:

$$-\frac{h_k^2}{2}\frac{f''(a_k+\theta h_k)}{f'(a_k)}=h_k+\frac{f(a_k)}{f'(a_k)}=\alpha-\left(a_k-\frac{f(a_k)}{f'(a_k)}\right)=\alpha-a_{k+1}=h_{k+1}.$$

De aquí

$$|h_{k+1}|=h_k^2\left|\frac{f''(a_k+\theta h_k)}{2f'(a_k)}\right|<h_k^2\frac{B}{2A}=Ch_k^2, \qquad k=0,\ 1,\ 2,\ \ldots$$

Por lo tanto,

$$|h_{k+1}|<Ch_k^2<C^3h_{k-1}^4<C^7h_{k-2}^8<\ldots<C^{2^{k+1}-1}h_0^{2^{k+1}},$$

o bien, como $|h_0|=|\alpha-a_0|<b-a$,

$$|h_{k+1}|<C^{-1}[C(b-a)]^{2^{k+1}}, \qquad k=0,\ 1,\ 2,\ \ldots \qquad (7)$$

En virtud de la condición (4), de aquí se deduce que *la diferencia h_k entre la raíz α y su valor aproximado a_k, obtenido por aplicación reiterada del método de Newton, tiende a cero al crecer k*, como se quería demostrar.

Señalemos que la fórmula (7) da una *cota del error* para la $(k+1)$-ésima aproximación, lo cual es importante si el método de Newton se aplica solo, y no en combinación con el método de interpolación lineal.

En los cursos de la teoría del cálculo aproximado, el lector podrá hallar procedimientos más racionales para realizar los cálculos con los métodos expuestos. En estos mismos cursos se puede hallar la exposición de muchos métodos de cálculo aproximado de raíces. *Entre éstos, el más perfecto es el método de Lobachevski* (a veces, llamado equivocadamente método de Gräffe). Este método permite hallar simultáneamente los valores aproximados de todas las raíces, incluyendo las imaginarias, sin exigir la separación previa de ellas; no obstante, requiere cálculos muy complicados. Este método se basa *en la teoría de los polinomios simétricos expuesta en el cap.* 11.

CAPITULO X

CAMPOS Y POLINOMIOS

§ 43. Anillos y campos numéricos

En muchos de los apartados anteriores del curso nos encontrábamos en la situación siguiente: exponiendo un tema, se permitía operar, o bien con números complejos arbitrarios, o bien solamente con números reales. Pero, después advertimos que los resultados obtenidos tienen también valor cuando se consideran solamente números reales (o que se generalizan respectivamente, palabra por palabra, para el caso de números complejos arbitrarios). Por regla general, en todos estos casos se podía observar que la teoría expuesta se conservaría enteramente también en el caso en que se permitiese tratar solamente con números racionales. Ha llegado ya el momento de explicar al lector las causas verdaderas de este paralelismo, para exponer el material ulterior en su generalidad natural, es decir, en el idioma algebraico usual. Con este fin, introduciremos el concepto de **campo** y también el de **anillo**. A pesar de ser este último un concepto más amplio, en nuestro curso va a desempeñar un papel auxiliar.

Está claro que los sistemas de todos los números complejos, de todos los números reales y de todos los números racionales, al igual que el sistema de todos los números enteros, *poseen la propiedad común de que en cada uno de ellos, manteniéndose dentro de los límites del mismo sistema, no sólo se puede efectuar la suma y el producto, sino también la resta.* Esta propiedad de los sistemas numéricos indicados los distingue, por ejemplo, del sistema de los números enteros positivos o de los números reales positivos.

Todo sistema de números complejos, o, en particular, reales, que contiene la suma, la diferencia y el producto de dos cualesquiera de sus números, se llama *anillo numérico.* Por lo tanto, los sistemas de todos los números enteros, racionales, reales o complejos, son anillos numéricos. Por otra parte, ningún sistema de números positivos será un anillo, pues, si a y b son dos números positivos diferentes, entonces, $a - b$ o $b - a$ será negativo. Un sistema cualquiera de números negativos tampoco será anillo, aunque sólo sea por el hecho de que el producto de dos números negativos es positivo.

Con los cuatro ejemplos considerados anteriormente no se agotan ni mucho menos los anillos numéricos. Ahora se van a señalar otros ejemplos, cuya comprobación de que el sistema considerado de números es verdaderamente anillo, se dejará al lector.

Los números pares forman un anillo; en general, para cualquier número natural n, el conjunto de números enteros divisibles por n es un anillo. Los números impares no forman anillo, pues la suma de dos números impares es par.

Es anillo el conjunto de los números racionales cuyos denominadores, en las expresiones en forma de fracciones irreducibles, son potencias del número 2; en particular, a este conjunto pertenecen todos los números enteros, pues, sus expresiones irreducibles tienen en el denominador el número 1, o sea, dos elevado a la potencia cero. En este ejemplo, en lugar del número 2 se podría tomar, naturalmente, cualquier número primo p. En general, tomando cualquier conjunto de números primos, finito e incluso infinito, y considerando el sistema de los números racionales cuyos denominadores, en las expresiones en forma de fracciones irreducibles, pueden dividirse solamente por los números primos que pertenecen al conjunto considerado, se obtiene también un anillo. Por otra parte, el conjunto de los números racionales cuyos denominadores, en las expresiones en fracciones irreducibles no se dividen por el cuadrado de ningún número primo, no es un anillo, puesto que la propiedad indicada no se conserva al multiplicar.

Veamos ejemplos de anillos numéricos que no pertenecen enteramente al anillo de los números racionales. El conjunto de los números de la forma

$$a+b\sqrt{2}, \tag{1}$$

donde a y b son números racionales arbitrarios, es un anillo; a éste pertenecen, en particulai, todos los números racionales (cuando $b=0$), y también el mismo número $\sqrt{2}$ (cuando $a=0$, $b=1$). También obtendríamos un anillo, si nos limitásemos a considerar solamente los números de la forma (1) con coeficientes enteros a, b. Claro, en estos ejemplos, en lugar del número $\sqrt{2}$ se podría tomar $\sqrt{3}$ o $\sqrt{5}$, etc.

El sistema de números de la forma

$$a+b\sqrt[3]{2} \tag{2}$$

con cualesquiera coeficientes racionales (o solamente con enteros cualesquiera) a, b, no forma anillo, pues, como fácilmente se com-

prueba, el producto del número $\sqrt[3]{2}$ por sí mismo no puede ser expresado en la forma (2) *.

Sin embargo, el sistema de números de la forma

$$a+b\sqrt[3]{2}+c\sqrt[3]{4} \tag{3}$$

con cualesquiera coeficientes racionales a, b, c, es ya un anillo, y esto mismo tiene lugar si se considera el caso de coeficientes enteros.

Examinemos ahora todos los números reales que se pueden obtener aplicando varias veces las operaciones de adición, multiplicación y sustracción al número π, bien conocido por el lector, y a números racionales cualesquiera. Estos son los números que se pueden escribir en la forma

$$a_0+a_1\pi+a_2\pi^2+\ldots+a_n\pi^n, \tag{4}$$

donde $a_0, a_1, a_2, \ldots, a_n$ son números racionales, $n \geqslant 0$. Obsérvese que ningún número puede poseer dos expresiones distintas de la forma (4), puesto que, en caso contrario, tomando la diferencia de dos expresiones de éstas obtendríamos que el número π tendría que satisfacer a una ecuación de coeficientes racionales; sin embargo, con los métodos del análisis matemático se demuestra que π no puede satisfacer a ninguna ecuación de coeficientes racionales, o sea, es un número trascendente. Por cierto, sin aplicar este resultado, o sea, sin suponer que la expresión de un número en la forma (4) sea única, se puede demostrar que los números de la forma (4) forman un anillo.

El conjunto de los números que se obtienen del número π y de los números racionales aplicando varias veces las operaciones de sumar, multiplicar, restar y dividir, es también anillo. Para la demostración no hay necesidad de buscar alguna expresión especial buena para los números considerados (a pesar de que podría hallarse): si los números

* En efecto, supongamos que

$$\sqrt[3]{4}=a+b\sqrt[3]{2}, \tag{2'}$$

donde los números a y b son racionales. Multiplicando ambos miembros de esta igualdad por $\sqrt[3]{2}$, obtenemos:

$$2=a\sqrt[3]{2}+b\sqrt[3]{4}.$$

Poniendo aquí la expresión (2′) para $\sqrt[3]{4}$, después de ciertas transformaciones simples llegamos a la igualdad

$$(a+b^2)\sqrt[3]{2}=2-ab. \tag{2''}$$

Si $a+b^2 \neq 0$, resulta,

$$\sqrt[3]{2}=\frac{2-ab}{a+b^2},$$

lo cual es imposible, pues, el segundo miembro es un número racional. Si $a+b^2=0$, en virtud de (2′′), también $2-ab=0$. De estas dos igualdades resulta que $b^3=-2$, lo cual de nuevo es imposible, pues el número b es racional.

α y β *se han obtenido* del número π y de ciertos números racionales, empleando las operaciones indicadas, esto mismo es cierto para los números $\alpha+\beta$, $\alpha-\beta$, $\alpha\beta$, y también (siendo $\beta \neq 0$) para el número $\frac{\alpha}{\beta}$.

Por fin, tomando el *conjunto de números complejos* $a+bi$ con cualesquiera coeficientes **racionales** a, b, se obtiene un anillo; esto mismo resulta cuando nos limitamos a coeficientes **enteros** a, b.

Los ejemplos examinados no pueden dar una idea completa de la diversidad de anillos numéricos existentes. Sin embargo, aquí *no vamos a continuar la lista de ejemplos y pasaremos a estudiar un* caso especial y muy importante de anillos numéricos. Por supuesto, ya sabemos que en los sistemas de todos los números racionales, de todos los números reales y de todos los números complejos, se puede efectuar la división ilimitadamente (excepto la división por cero), *mientras que la división de los números enteros sale fuera de los* límites del sistema de estos números. Hasta ahora no habíamos prestado atención a esta distinción pero, en realidad, es muy importante y conduce a la definición siguiente.

Un anillo numérico se llama *campo numérico*, si éste contiene el *cociente de dos cualesquiera de sus números (se supone que el divisor* es diferente de cero). Por consiguiente, se puede hablar del campo de números racionales, del campo de números reales, del campo de números complejos, por otra parte, el anillo de los números enteros no forma un campo.

El *realidad, algunos de los ejemplos considerados anteriormente* de anillos numéricos son campos. Ante todo, obsérvese que no existen campos numéricos distintos del campo de números racionales y contenidos totalmente en éste (no se considera campo el sistema formado por el cero único). Se cumple incluso la siguiente afirmación más general:

El campo de números racionales está contenido totalmente en cualquier campo numérico.

En efecto, sea dado un campo numérico, que designaremos con la letra P. Si a es un número arbitrario del campo P y diferente de *cero, P contiene también el cociente de la división del número a* por sí mismo, o sea, el número uno. Sumando unas cuantas veces la unidad consigo misma, obtenemos que todos los números naturales están contenidos en el campo P. Por otra parte, en el campo P tiene que estar contenida la diferencia $a-a$, o sea, el número cero, y, *por esto, también pertenece a P el resultado que se obtiene al restar* de cero cualquier número natural, es decir, cualquier número entero negativo. Finalmente, en el campo P están contenidos también los cocientes de los números enteros, o sea, en general, todos los números racionales.

En el campo de los números complejos están contenidos muchos campos distintos, siendo el campo de números racionales el menor de ellos. Así, pues, el anillo de los números de la forma

$$a+b\sqrt{2} \tag{5}$$

con coeficientes **racionales** (y no sólo con enteros) **arbitrarios** a, b, es un campo. En efecto, consideremos el cociente de dos números de la forma (5), $a+b\sqrt{2}$ y $c+d\sqrt{2}$, donde se supone que este último es diferente de cero; por consiguiente, también es diferente de cero el número $c-d\sqrt{2}$ de donde,

$$\frac{a+b\sqrt{2}}{c+d\sqrt{2}}=\frac{(a+b\sqrt{2})(c-d\sqrt{2})}{(c+d\sqrt{2})(c-d\sqrt{2})}=\frac{ac-2bd}{c^2-2d^2}+\frac{bc-ad}{c^2-2d^2}\sqrt{2}.$$

Hemos obtenido de nuevo un número de la forma (5), manteniéndose racionales los coeficientes. Naturalmente, en este ejemplo se puede sustituir el número $\sqrt{2}$ por la raíz cuadrada de cualquier número racional, cuya raíz cuadrada no pudiese ser extraída en el mismo campo de números racionales. Así, pues, los números de la forma $a+bi$ con coeficientes racionales a, b forman un campo.

§ 44. Anillo

En distintas ramas de las matemáticas y en sus aplicaciones, suele ocurrir frecuentemente que las operaciones algebraicas no se efectúan con números, sino con objetos de naturaleza distinta. En los capítulos anteriores se pueden hallar muchos de estos ejemplos recordemos el producto y la suma de matrices, la suma de vectores, las operaciones con los polinomios, las operaciones con las transformaciones lineales. La definición general de *operación algebraica* a que satisfacen las operaciones de sumar y de multiplicar en los anillos numéricos, y también las operaciones en los ejemplos indicados, consiste en lo siguiente.

Sea dado un conjunto M que conste de números o de objetos de naturaleza geométrica, o en general, de algunos entes, que llamaremos *elementos* de este conjunto. Se dice que *en el conjunto M está definida una operación algebraica*, si está indicada una regla según la cual, **a cada par** de elementos a, b de este conjunto se pone en correspondencia **de un modo unívoco** un tercer elemento c, perteneciente también a M. Esta operación puede llamarse adición (o suma), y entonces, c se llamará *suma* de los elementos a y b, representándose con la notación $c=a+b$; esta operación puede llamarse *multiplicación*, o sea, c será el *producto* de los elementos a y b ($c=ab$); finalmente, es posible que para la operación definida en el conjunto M se introduzca una nueva terminología y simbolismo.

En cada uno de los anillos numéricos están definidas dos operaciones independientes, la adición y la multiplicación. En lo que se refiere a la resta y a la división, éstas no pueden considerarse operaciones nuevas, pues son las *inversas* de la adición y *multiplicación*, respectivamente, si convenimos en tomar la siguiente definición general de *operación inversa*.

Supongamos que en el conjunto M está definida una operación algebraica, por ejemplo, la suma. Se dice que *para esta operación existe una operación inversa*, la resta, si **para cada** par de elementos a, b de M, existe en M un elemento d, **unívocamente determinado**, que satisface a la igualdad: $b + d = a$. En este caso, el elemento d *se llama diferencia de los elementos* a *y* b y *se designa con la notación* $d = a - b$.

Está claro que tanto la suma como la multiplicación poseen operación inversa en los campos numéricos (por cierto, la multiplicación *con cierta restricción: el divisor tiene que ser diferente de* cero). En los anillos numéricos que no son campos (como, por ejemplo, en el anillo de los números enteros), solamente la suma posee operación inversa.

Por otra parte, en el sistema de todos los polinomios en la indeterminada x, cuyos coeficientes pertenecen a un campo numérico fijado P, también están definidas dos operaciones: la suma y el producto; además, *la suma posee operación inversa: la resta*.

Como se sabe, tanto en los anillos numéricos como en el sistema de los polinomios, las operaciones de sumar y multiplicar poseen las propiedades siguientes (a, b, c, son números arbitrarios del anillo *numérico considerado o polinomios arbitrarios del sistema considerado*):

I. La adición es conmutativa: $a + b = b + a$.

II. La adición es asociativa: $a + (b + c) = (a + b) + c$.

III. La *multiplicación es conmutativa:* $ab = ba$.

IV. La multiplicación es asociativa: $a\,(bc) = (ab)\,c$.

V. La adición y la multiplicación están ligadas por la ley distributiva:

$$(a + b)\,c = ac + bc.$$

Ahora ya estamos preparados para hacer la definición general del concepto de anillo, que es uno de los conceptos fundamentales del álgebra.

Un conjunto R se denomina *anillo*, si se han definido en él dos operaciones, llamadas adición o suma y multiplicación, siendo ambas *conmutativas* y *asociativas*, y *ligadas por la ley distributiva*, poseyendo además la suma la operación inversa, llamada resta.

Por lo tanto, son ejemplos de anillos, los anillos numéricos y los anillos de polinomios en la indeterminada x con coeficientes de un

campo numérico dado e incluso de un anillo numérico dado. Señalemos otro ejemplo más que aclara con amplitud el concepto de anillo.

El curso de análisis matemático comienza con la definición de **función** de la variable real x. Consideremos el conjunto de las funciones, determinadas para **todos** los valores reales de x y que toman valores reales. En él definiremos las operaciones algebraicas del modo siguiente: la *suma* de dos funciones $f(x)$ y $g(x)$ será una función cuyo valor para cualquier $x = x_0$ será igual a la suma de los valores de las funciones dadas, o sea, igual a $f(x_0) + g(x_0)$; el *producto* de estas funciones será una función cuyo valor para cualquier $x = x_0$ será igual al producto $f(x_0) \cdot g(x_0)$. Es evidente que la suma y el producto existen para cualesquiera dos funciones del conjunto considerado. La validez de las propiedades I-V se comprueba sin dificultad alguna: la suma y multiplicación de funciones se reducen a la suma y multiplicación de sus valores para cualquier x, es decir, a operaciones con números reales para los que se cumplen las propiedades I-V. Finalmente, tomando por *diferencia* de las funciones $f(x)$ y $g(x)$ la función cuyo valor para cualquier $x = x_0$ sea igual a la diferencia $f(x_0) - g(x_0)$, obtenemos la sustracción, operación inversa a la adición. Con esto queda demostrado que *el conjunto de las funciones determinadas para todas las x reales, después de haber introducido del modo descrito anteriormente las operaciones de sumar y multiplicar, se convierte en un anillo.*

Se pueden obtener otros ejemplos de anillos de funciones, conservando las definiciones de las operaciones con las funciones dadas anteriomente, pero, considerando, por ejemplo, las funciones determinadas sólo para los valores positivos de la variable x o las funciones determinadas para los valores x del segmento $[0, 1]$. En general, el sistema de todas las funciones que tienen un campo dado de definición, es un anillo. También se podrían obtener ejemplos de anillos sin considerar todas las funciones determinadas en un campo dado, sino solamente las funciones continuas que se estudian en el curso de análisis matemático. Por otro lado, se podrían considerar las funciones complejas de variable compleja. Existen muchísimos anillos distintos de funciones, así como distintos anillos numéricos.

Establezcamos algunas **propiedades elementales** de los anillos que se deducen inmediatamente de su definición.

Estas propiedades son ordinarias para el caso de los números, sin embargo, pueda ser que al lector le sorprenda que éstas sean consecuencia solamente de las condiciones I-V y de la existencia unívoca de la resta.

Hagamos primero unas cuantas observaciones sobre la importancia de las condiciones I—V. El papel de las *leyes conmutativas* no necesita explicaciones. El valor de las *leyes asociativas* consiste

en lo siguiente: *en la definición de la operación algebraica se trata de la suma o del producto de dos elementos solamente*. Si, por ejemplo, probamos definir el producto de tres elementos a, b, c, nos *encontramos con* la dificultad de que, por lo general, los productos au y vc, donde $bc = u$, $ab = v$, pueden no coincidir, o sea, $a(bc) \neq (ab)c$. La ley asociativa exige que estos productos sean iguales a un mismo elemento del anillo: resulta natural tomar este elemento por producto abc, escribiéndolo ya sin paréntesis. *La ley asociativa permite también definir unívocamente el producto (respectivamente, la suma) de cualquier número finito de elementos del anillo*, es decir, permite demostrar la independencia del producto de cualesquiera n elementos de la distribución primaria de los paréntesis.

Demostremos esta afirmación por el método de inducción sobre n. Esta se ha demostrado ya para $n = 3$, por lo cual suponemos que $n > 3$ y que nuestra afirmación ya está demostrada para todos los números menores que n. Sean dados los elementos $a_1, a_2, \ldots, a_n$ y supongamos que en este sistema se han *distribuido los paréntesis de algún modo, indicando el orden en que se debe efectuar la multiplicación*. La última operación consistirá en multiplicar el producto de los primeros k elementos $a_1a_2 \ldots a_k$ (donde $1 \leqslant k \leqslant n - 1$) por el producto $a_{k+1}\, a_{k+2} \ldots a_n$. Como estos productos constan de un número de factores menor *que* n, *y que por la hipótesis están definidos unívocamente, no queda más que* demostrar la igualdad

$$(a_1a_2 \ldots a_k)(a_{k+1}a_{k+2} \ldots a_n) = (a_1a_2 \ldots a_l)(a_{l+1}a_{l+2} \ldots a_n).$$

para cualesquiera k y l.

Con este fin, es suficiente considerar el caso $l = k + 1$. En este caso, poniendo

$$a_1a_2 \ldots a_k = b, \quad a_{k+2}\, a_{k+3} \ldots a_n = c,$$

y, basándonos en la ley asociativa, obtenemos

$$b(a_{k+1}c) = (ba_{k+1})c.$$

Con esto queda demostrada nuestra afirmación.

En particular, se puede hablar del producto de n elementos *iguales entre sí, o sea, se puede introducir el concepto de potencia* a^n del elemento a con exponente entero y positivo n. Se comprueba fácilmente que son válidos en cualquier anillo las reglas de operación con los exponentes. De modo análogo, la ley asociativa de la adición nos lleva al concepto de *múltiplo* na del elemento a con un coeficiente entero y positivo n.

La *ley distributiva*, es decir, la regla ordinaria para abrir paréntesis, es la única exigencia en la definición de anillo que liga la suma y la multiplicación; el hecho de que el estudio simultáneo de las dos operaciones indicadas proporcione algo más que lo que se podría obtener al estudiarlas por separado, se debe solamente a esta ley. En la formulación de la ley distributiva participa únicamente la suma de dos términos. Pero sin dificultad se demuestra

que se verifica la igualdad

$$(a_1 + a_2 + \ldots + a_k)\, b = a_1 b + a_2 b + \ldots + a_k b$$

para cualquier k, y la regla general para multiplicar una suma por otra.

En cualquier anillo también se cumple la ley distributiva para la resta. En efecto, según la definición de la resta, el elemento $a - b$ satisface a la igualdad

$$b + (a - b) = a.$$

Multiplicando por c ambos miembros de esta igualdad y aplicando al primer miembro de ésta la ley distributiva, obtenemos:

$$bc + (a - b)\, c = ac.$$

Por consiguiente, el elemento $(a - b)\, c$ es la diferencia de los elementos ac y bc:

$$(a - b)\, c = ac - bc.$$

De la existencia de la resta se deducen unas propiedades muy importantes de los anillos. Si a es un elemento arbitrario del anillo R, la diferencia $a - a$ es un elemento del anillo completamente determinado. Su papel es análogo al del cero en los anillos numéricos, mas, según la definición, éste puede depender de la elección del elemento a y, por esto, lo designaremos por ahora mediante 0_a.

Demostremos que para todos los a, los elementos 0_a son iguales entre sí. En efecto, si b es otro elemento arbitrario del anillo R, agregando el elemento 0_a a ambos miembros de la igualdad

$$a + (b - a) = b$$

y aplicando la igualdad $0_a + a = a$, resulta:

$$0_a + b = 0_a + a + (b - a) = a + (b - a) = b.$$

Por lo tanto,

$$0_a = b - b = 0_b.$$

Hemos demostrado que *todo anillo R posee un elemento unívocamente determinado, cuya suma con cualquier elemento a de este anillo es igual a a.* Este elemento se llamará cero del anillo R y se designará con el símbolo 0, no representando un peligro serio el que sea confundido con el número cero. Por lo tanto,

$$a + 0 = a \text{ para todos los elementos } a \text{ de } R$$

En cualquier anillo, para cada elemento a existe un elemento opuesto $-a$ unívocamente determinado que satisface a la igualdad:

$$a + (-a) = 0;$$

precisamente, este elemento es la diferencia $0-a$; la unicidad es consecuencia de la unicidad de la resta. Evidentemente $-(-a)=a$. La diferencia $b-a$ de dos elementos cualesquiera del anillo se puede escribir ahora de la forma

$$b-a=b+(-a).$$

En efecto,

$$[b+(-a)]+a=b+[(-a)+a]=b+0=b.$$

Para cualquier elemento a de un anillo y cualquier número entero positivo n, se cumple la igualdad:

$$n(-a)=-(na).$$

En efecto, agrupando los términos resulta:

$$na+n(-a)=n[a+(-a)]=n\cdot 0=0.$$

Hemos obtenido ahora la posibilidad de definir los *múltiplos negativos* de un elemento del anillo: siendo $n>0$, los elementos iguales $n(-a)$ y $-(na)$ se designarán mediante $(-n)a$. Finalmente, convengamos tomar por cero del anillo considerado el *múltiplo nulo* $0\cdot a$ de cualquier elemento.

Hemos dado la definición del cero empleando solamente la suma y su operación inversa, o sea, sin utilizar la multiplicación. Sin embargo, en el caso de los números, el cero posee respecto a la multiplicación una propiedad característica, que es además muy importante. El cero de cualquier anillo posee la propiedad: *en cualquier anillo, el producto de cualquier elemento por el cero es igual a cero.* La demostración se basa directamente en la ley distributiva: siendo a un elemento arbitrario del anillo R, cualquiera que sea el elemento auxiliar x de este anillo, se tiene:

$$a\cdot 0=a(x-x)=ax-ax=0.$$

Aplicando esta propiedad del cero se puede demostrar que *en cada anillo, para cualesquiera elementos a, b, se cumple la igualdad:*

$$(-a)b=-ab.$$

En efecto,

$$ab+(-a)b=[a+(-a)]b=0\cdot b=0.$$

De aquí se deduce que la conocida regla de la multiplicación de los números negativos, «menos por menos da más», también se deduce de la definición de anillo, es decir, que *en cualquier anillo se verifica la igualdad*

$$(-a)(-b)=ab.$$

En efecto,

$$(-a)(-b)=-[a(-b)]=-(-ab)=ab.$$

El lector demostrará ahora sin dificultad que en cualquier anillo, para los múltiplos (incluyendo los negativos) de cualquier elemento son válidas todas las reglas de operaciones con los múltiplos de un número.

Por lo tanto, las operaciones algebraicas en cualquier anillo poseen muchas propiedades ordinarias de las operaciones con los números. Sin embargo, no hay que creer que en cualquier anillo se conservan todas las propiedades de la suma y multiplicación de los números. Así, pues, la multiplicación de los números posee una propiedad que es recíproca a la considerada anteriormente: **si el producto de dos números es igual a cero, al menos uno de los factores es igual a cero.** Esta propiedad ya no se puede generalizar para cualquier anillo, pues, en algunos anillos se pueden señalar pares de elementos diferentes de cero, cuyo producto es igual a cero, es decir, $a \neq 0$, $b \neq 0$, pero $ab = 0$; los elementos a, b, que poseen esta propiedad se llaman *divisores de cero*.

Claro, entre los anillos numéricos no se pueden hallar ejemplos de anillos con divisores de cero. Tampoco contienen divisores de cero los anillos de polinomios de coeficientes numéricos. Pero hay muchos anillos de funciones que poseen divisores de cero. Obsérvese primeramente que en cualquier anillo de funciones el cero es la función que se convierte en cero para todos los valores de la variable x. Consideremos ahora las funciones $f(x)$ y $g(x)$ que siguen, definidas para todos los valores reales de x:

$$f(x) = 0 \text{ para } x \leqslant 0, \qquad f(x) = x \text{ para } x > 0;$$
$$g(x) = x \text{ para } x \leqslant 0, \qquad g(x) = 0 \text{ para } x > 0.$$

Estas funciones son diferentes de cero, pues, sus valores no son iguales a cero para todos los valores de x; sin embargo, el producto de estas funciones es igual a cero.

No todas las condiciones I-V que figuran en la definición de anillo son necesarias en igual medida. El desarrollo de la ciencia muestra que mientras las propiedades I y II de la suma y la ley distributiva V se cumplen en todas las aplicaciones, la introducción de las propiedades III y IV de la multiplicación en la definición de anillo resulta demasiado incómoda, reduciendo el posible campo de aplicación de este concepto. Así, pues, el conjunto de las matrices cuadradas de orden n de elementos reales, considerado con las operaciones de adición y multiplicación de matrices, satisface a todas las condiciones que figuran en la definición de anillo, excluyendo la ley conmutativa de la multiplicación. Las multiplicaciones no conmutativas aparecen con tanta frecuencia y en casos tan importantes que actualmente el término de «anillo» se entiende ordinariamente como anillo *no conmutativo* (mejor dicho, como un anillo que no es necesariamente conmutativo, en el sentido de que la multiplicación puede ser no conmutativa), llamando *anillo conmutativo* al tipo particular de anillos en los que se cumple la condición III.

Ultimamente ha aumentado el interés hacia los anillos con multiplicación no asociativa, elaborándose ya la teoría general de los anillos como la teoría

de los anillos no asociativos (es decir, que no son necesariamente asociativos). El conjunto de vectores del espacio euclídeo de tres dimensiones respecto a las operaciones de la suma y de la multiplicación vectorial (conocida por el curso de geometría analítica) es un ejemplo simple de tales anillos.

§ 45. Campo

Del mismo modo que entre los anillos numéricos fueron elegidos y denominados campos numéricos aquellos anillos en los que se podía efectuar la división (excepto la división por cero), resulta natural hacer lo mismo en el caso general. Obsérvese primeramente que *en ningún anillo es posible la división por cero*, debido a la propiedad del cero respecto a la multiplicación, demostrada anteriormente: dividir un elemento a por cero significa hallar en el anillo un elemento x tal, que $0 \cdot x = a$, lo cual es imposible si $a \neq 0$, pues, el primer miembro es igual a cero.

Hagamos la definición siguiente:

Un anillo P se llama *campo*, si consta no sólo del cero y en él es posible la división en todos los casos (a excepción de la división por cero), determinándose ésta unívocamente, o sea, si para cualesquiera elementos a, b de P, **de los cuales** b **es diferente de cero**. existe en P un elemento q, y sólo uno, que satisface a la igualdad: $bq = a$. El elemento q se llama *cociente* de los elementos a y b y se designa con la notación $q = \frac{a}{b}$ *.

Naturalmente, todos los campos numéricos son ejemplos de campos. El anillo de los polinomios en la indeterminada x con coeficientes reales o, en general, con coeficientes de algún campo numérico, no es campo: la división con resto que existe para los polinomios se diferencia, naturalmente, de la división «exacta», supuesta en la definición de campo. Por otra parte, se ve fácilmente que *el conjunto*

* En realidad, la unicidad de la división en un campo, así como la unicidad de la resta, supuesta en la definición de anillo, se puede demostrar sin dificultad aplicando otras condiciones que figuran en la definición de campo o, respectivamente, de anillo (*Nota del A.*).

Un caso más general resulta cuando no se insiste en que la operación de multiplicar satisfaga a la ley conmutativa (o sea, cuando el anillo puede ser no conmutativo; véase la última parte del § 44). En este caso, además del elemento q, tiene que existir en P un elemento q' (que puede ser distinto de q), y sólo uno, que satisfaga a la igualdad: $q'b = a$. El anillo P se llama entonces cuerpo. Por lo tanto, se puede decir que campo es un cuerpo conmutativo.

Según parece, el vocablo «campo», para la denominación abreviada de un cuerpo conmutativo, ha sido empleado por primera vez en castellano por R. Rodríguez Vidal, en su traducción de la obra de Birkhoff y MacLane «Algebra Moderna». Teniendo también en cuenta que en los libros soviéticos, el vocablo «поле» («campo») está ya admitido hace muchos años, creemos conveniente emplear a continuación este último como traducción del primero. (*Nota del T.*).

de las funciones racionales con coeficientes reales (véase el § 25) *forma un campo* que contiene al anillo de los polinomios, del mismo modo que el campo de los números racionales contiene al anillo de los números enteros.

Entre los anillos de funciones se pueden indicar otros ejemplos de campos; sin embargo, aquí no vamos a detenernos en ellos y pasaremos a examinar otros ejemplos de distinto género.

Todos los anillos numéricos y, en general, los anillos que hasta ahora hemos examinado, contienen una infinidad de elementos. Sin embargo, existen anillos e incluso campos que constan de un número finito de elementos. Los ejemplos más simples de *anillos finitos* y *campos finitos*, empleados esencialmente en la teoría de los números, se forman del modo siguiente.

Se toma un número natural cualquiera n, diferente de 1. Los números enteros a y b se llaman *congruentes respecto del módulo* n,

$$a \equiv b \,(\text{mód}\, n),$$

si al dividirlos por n dan un mismo residuo, o sea, si su diferencia es divisible por n. Todo el anillo de los números enteros se descompone en n clases disjuntas,

$$C_0,\ C_1,\ \ldots,\ C_{n-1}, \tag{1}$$

de números congruentes entre sí respecto del módulo n, donde la clase C_k, $k = 0, 1, \ldots, n-1$, consta de los números que al dividirlos por n dan el residuo k. Resulta que se puede definir la suma y el producto de estas clases de un modo muy natural.

Con este fin, tomemos unas clases cualesquiera C_k y C_l (no necesariamente distintas) del sistema (1). Sumando cualquier número de la clase C_k con cualquier número de la clase C_l, obtenemos cada vez números que pertenecen a una clase determinada: a la clase C_{k+l}, si $k+l<n$, o a la clase C_{k+l-n}, si $k+l \geqslant n$. Esto nos lleva a la siguiente definición de *suma de las clases*:

$$\begin{aligned} C_k + C_l &= C_{k+l} \quad \text{si} \quad k+l<n, \\ C_k + C_l &= C_{k+l-n} \quad \text{si} \quad k+l \geqslant n. \end{aligned} \tag{2}$$

Por otra parte, multiplicando cualquier número de la clase C_k por cualquier número de la clase C_l, obtenemos números que están de nuevo en una clase determinada: precisamente en la clase C_r, donde r es el residuo de la división del producto kl por n. Por lo tanto, tomamos la definición siguiente de *producto de clases*:

$$C_k \cdot C_l = C_r, \text{ donde } kl = nq + r,\ 0 \leqslant r < n. \tag{3}$$

El sistema (1) *de clases de números enteros, congruentes entre sí respecto del módulo* n, *es un anillo respecto de las operaciones definidas*

por las condiciones (2) *y* (3). En efecto, la validez de las condiciones I-V de la definición de anillo se establece comprobándolas directamente. Además, es también consecuencia de la validez de estas *condiciones en el anillo de los números enteros y de la relación* indicada anteriormente entre las operaciones con los números enteros y las operaciones con las clases. Está claro que la clase C_0, compuesta de los números divisibles por n, desempeña el papel del cero. El elemento opuesto para la clase C_k, $k = 1, 2, \ldots, n-1$, es la clase C_{n-k}. Por consiguiente, en el sistema de las clases (1) se puede definir la resta, es decir, este sistema satisface a todas las condiciones que figuran en la definición de *anillo*. Convengamos en designar el anillo obtenido mediante Z_n.

Si el número n es compuesto, el anillo Z_n posee divisores de cero y, por esto, como se demostrará más abajo, no puede ser campo. En efecto, si $n = kl$, donde $1 < k < n$, $1 < l < n$, las clases C_k y C_l son distintas de la clase cero C_0, pero, según la definición del producto de las clases (véase (3)), $C_k \cdot C_l = C_0$.

Si el número n es primo, el anillo Z_n es un campo.

En efecto, sean dadas las clases C_k y C_m, donde $C_k \neq C_0$, o sea, $1 \leqslant k \leqslant n-1$. Hay que demostrar que se puede dividir C_m por C_k, o sea, que se puede hallar una clase C_l tal, que $C_k \cdot C_l = C_m$. Si $C_m = C_0$, se tiene $C_l = C_0$. Si $C_m \neq C_0$ consideramos el sistema de números

$$k, \ 2k, \ 3k, \ \ldots, \ (n-1)\,k. \tag{4}$$

Todos estos números están fuera de la clase cero C_0, pues, el producto de dos números naturales menores que el número primo n no puede ser divisible por éste. Por otra parte, *ninguno* de los dos números sk y tk del sistema (4), $s < t$, puede estar situado en una clase, puesto que, en caso contrario, su diferencia

$$tk - sk = (t - s)\,k$$

sería divisible por n, lo cual es absurdo, debido a que el número n es primo. Por lo tanto, en cada clase no nula está situado exactamente *un número del sistema* (4). *En particular, en la clase C_m está situado* el número lk, donde $1 \leqslant l \leqslant n-1$, o sea, $C_l \cdot C_k = C_m$, y entonces la clase C_l es el cociente buscado de la división de C_m por C_k.

Por consiguiente, hemos obtenido una infinidad de campos finitos distintos: el campo Z_2, compuesto de dos elementos solamente, y también los campos Z_3, Z_5, Z_7, Z_{11}, etc.

Ahora veremos **algunas propiedades de los campos** que se deducen *de la existencia de la división. Estas propiedades son análogas* a las propiedades de los anillos basados en la existencia de la resta y se demuestran con los mismos razonamientos, por lo cual, la demostración la dejamos al lector.

Todo campo P posee un elemento, unívocamente determinado, cuyo producto por cualquier elemento a de este campo es igual a a. Este elemento, que coincide con los cocientes iguales entre sí $\frac{a}{a}$ para todos los a, diferentes de cero, se llama *unidad* del campo P y se designa con el símbolo 1. Por lo tanto,

$$a \cdot 1 = a \text{ para todos los elementos } a \text{ de } P.$$

En todo campo, para cualquier elemento a diferente de cero, existe un elemento recíproco a^{-1}, unívocamente determinado, que satisface a la igualdad

$$a \cdot a^{-1} = 1;$$

este elemento es precisamente $a^{-1} = \frac{1}{a}$. Está claro que $(a^{-1})^{-1} = a$. El cociente $\frac{b}{a}$ se puede escribir ahora en la forma

$$\frac{b}{a} = b \cdot a^{-1}.$$

Para cualquier elemento a diferente de cero, y cualquier entero positivo n, se verifica la igualdad

$$(a^{-1})^n = (a^n)^{-1}.$$

Designando estos elementos iguales entre sí mediante a^{-n}, obtenemos las *potencias negativas* de un elemento del campo, para las que rigen las reglas de operación ordinarias. Hagamos, finalmente, $a^0 = 1$ para todos los a.

La existencia de unidad no es una propiedad característica de los campos, pues, por ejemplo, el anillo de los números enteros posee unidad. Sin embargo, el ejemplo del anillo de los números pares muestra que no todos los anillos poseen unidad. Por otra parte, *todo anillo que posea unidad y que contenga al elemento recíproco de cualquier elemento diferente de cero, es un campo.* En efecto, en este caso el producto ba^{-1}, $a \neq 0$, servirá de cociente $\frac{b}{a}$. La unicidad de este cociente se demuestra sin dificultad alguna.

Obsérvese que *ningún campo contiene divisores de cero.* En efecto, sea $ab = 0$, pero $a \neq 0$. Multiplicando ambos miembros de la igualdad por el elemento a^{-1}, en el primer miembro resulta $(a^{-1}\,a)\,b = 1 \cdot b = b$, y en el segundo, $a^{-1} \cdot 0 = 0$, o sea, $b = 0$. De aquí se deduce que *en todo campo cualquier igualdad se puede simplificar por un factor común diferente de cero.* En efecto, si $ac = bc$ y $c \neq 0$, se tiene $(a - b)\,c = 0$, de donde $a - b = 0$, o sea, $a = b$.

De la definición del cociente $\frac{a}{b}$ (donde $b \neq 0$) y de la posibilidad, anteriormente demostrada, de escribirlo en forma de producto ab^{-1}, se puede demostrar sin dificultad que *en todo campo se conservan las reglas ordinarias de operación con los quebrados.* A saber,

$$\frac{a}{b} = \frac{c}{d} \text{ cuando, y sólo cuando, } ad = bc;$$

$$\frac{a}{b} \pm \frac{c}{d} = \frac{ad \pm bc}{bd};$$

$$\frac{a}{b} \cdot \frac{c}{d} = \frac{ac}{bd};$$

$$\frac{-a}{b} = -\frac{a}{b}.$$

Característica de un campo. No todas propiedades de los campos numéricos se conservan en el caso de un campo arbitrario. Así, pues, sumando así mismo el número 1 unas cuantas veces, o sea, tomando cualquier entero positivo que sea múltiplo de la unidad, nunca se obtendrá el cero. En general, todos estos múltiplos, es decir, todos los números naturales, son distintos entre sí. Si se toman enteros múltiplos de 1 de algún campo finito, entre ellos habrá indispensablemente algunos que sean iguales, pues este campo tiene sólo un número finito de elementos distintos. Si todos los múltiplos enteros de la unidad del campo P son elementos distintos de este campo, o sea, si $k \cdot 1 \neq l \cdot 1$ cuando $k \neq l$, se dice que P es un campo de *característica cero*; tales son, por ejemplo, todos los campos numéricos. Si existen unos números enteros k y l, $k > l$ tales, que en P se cumple la igualdad $k \cdot 1 = l \cdot 1$, entonces $(k - l) \cdot 1 = 0$, es decir, existe en P un múltiplo positivo de la unidad igual a cero, llamándose entonces *P campo de característica finita*. Precisamente ésta es igual a p, si p es el primer coeficiente positivo con el que se anula la unidad del campo P. Todos los campos finitos son ejemplos de campos de característica finita; existen también campos infinitos de característica finita.

Si p es la característica del campo P, el número p es primo.

En efecto, de la igualdad $p = st$, donde $s < p$, $t < p$, resultaría la igualdad $(s \cdot 1)(t \cdot 1) = p \cdot 1 = 0$, y como el campo no puede tener divisores de cero, se tendría $s \cdot 1 = 0$, o bien, $t \cdot 1 = 0$, lo cual contradice a la definición de la característica como el coeficiente positivo **menor** que convierte en cero a la unidad del campo.

Si la característica del campo P es igual a p, para cualquier elemento a de este campo se verifica la igualdad $pa = 0$. Si la característica del campo P es igual a cero, a es un elemento de este campo y n es un número entero, entonces las condiciones $a \neq 0$ y $n \neq 0$ implican la desigualdad $na \neq 0$.

En efecto, en el primer caso, el elemento pa, o sea, la suma de p términos iguales a a, sacando a fuera de paréntesis, se puede representar en la forma

$$pa = a\,(p\cdot 1) = a\cdot 0 = 0.$$

En el segundo caso, para $a \neq 0$, de la igualdad $na = 0$, o sea, $a\,(n\cdot 1) = 0$, resultaría la igualdad $n\cdot 1 = 0$, y, como la característica del campo es igual a cero, se tendría $n = 0$.

Subcampos, ampliaciones (extensiones). Supongamos que una parte de los elementos de un campo P, formando un conjunto P', también forma un campo con respecto a las operaciones definidas en el campo P, es decir, que para dos elementos cualesquiera a, b de P', los elementos $a + b$, ab, $a - b$, y para $b \neq 0$, $\frac{a}{b}$, contenidos en el campo P, también pertenecen a P' (claro, cumpliéndose las leyes I-V en P también se cumplen en P'). En este caso, P' se llama *subcampo* del campo P, y P, *ampliación* (o extensión) del campo P'. Es evidente que el cero y la unidad del campo P también están contenidos en P' y en éste sirven también de cero y unidad. Así, pues, el campo de los números racionales es un subcampo del campo de los números reales; todos los campos numéricos son subcampos del campo de los números complejos.

Supongamos que en el campo P se han dado un subcampo P' y un elemento c situado fuera de P', y que hemos hallado el subcampo mínimo P'' del campo P que contiene a P' y a c. Este subcampo mínimo tiene que ser único, pues si P''' fuese otro subcampo más con estas propiedades, la intersección de los subcampos P'' y P''' (o sea, el conjunto de los elementos comunes a ambos subcampos) contendría a P' y al elemento c y, junto con dos elementos cualesquiera suyos, contendría también a su suma (esta suma tiene que estar contenida en P'' y en P''' y, por lo tanto, en su intersección), y también a su producto, resta y cociente; en otras palabras, esta intersección misma sería un subcampo, lo cual es absurdo, pues el subcampo P'' es mínimo. Se dice que el campo P'' *se ha obtenido por adjunción del elemento c al campo P'*, empleándose la notación $P'' = P'(c)$.

Evidentemente, el campo $P'(c)$, además del elemento c y de todos los elementos del campo P', contiene también todos los elementos que se obtienen de ellos mediante la suma, multiplicación, resta y división. Como ejemplo, señalemos la ampliación del campo de los números racionales, considerado en el § 43, que consta de los números de la forma $a + b\sqrt{2}$ con racionales a, b; esta ampliación se obtiene por adjunción del número $\sqrt{2}$ al campo de los números racionales.

§ 46. Isomorfismo de los anillos (de los campos). Unicidad del campo de los números complejos

En la teoría de los anillos desempeña un gran papel el concepto de isomorfismo. Los anillos L y L' se llaman *isomorfos* si entre sus elementos se puede establecer una correspondencia biunívoca tal que, para cualesquiera elementos a, b de L y sus correspondientes elementos a', b' de L', a la suma $a + b$ le corresponda la suma $a' + b'$ y al producto ab, el producto $a'b'$.

Supongamos que entre los anillos L y L' se ha establecido una correspondencia de isomorfismo. *Entonces al cero* 0 *del anillo* L *le corresponde el cero* $0'$ *del anillo* L'. En efecto, supongamos que al elemento 0 le corresponde el elemento c' de L'. Tomemos un elemento arbitrario a de L y el elemento a' de L' que le corresponde. Entonces, al elemento $a + 0$ le tiene que corresponder el elemento $a' + c'$; pero como $a + 0 = a$, se tiene, $a' + c' = a'$, de donde $c' = 0'$. *Al elemento* $-a$ *le corresponde el elemento* $-a'$. En efecto, supongamos que al elemento $-a$ le corresponde el elemento d'. Entonces al elemento $a + (-a) = 0$ le tiene que corresponder el elemento $a' + d'$, o sea, $a' + d' = 0'$, de donde $d' = -a'$. De aquí resulta que *a la diferencia de elementos de* L *le corresponde la diferencia de los elementos correspondientes de* L'. Con razonamientos análogos se puede demostrar que, si el anillo L posee unidad, la imagen de este elemento (o sea, el elemento que le corresponde en L', en el isomorfismo considerado) es la unidad del anillo L', y si el elemento a de L tiene elemento recíproco a^{-1}, la imagen del elemento a^{-1} en L' es el elemento recíproco de a'.

De aquí se deduce que *un anillo que es isomorfo a un campo, es también un campo*. Fácilmente se ve también que la propiedad de un anillo de no tener divisores de cero se conserva también en la correspondencia de isomorfismo. En general, los anillos isomorfos pueden diferenciarse entre sí por la naturaleza de sus elementos, pero, por sus propiedades algebraicas, son idénticos. Cualquier teorema demostrado para un anillo subsiste también para los anillos que son isomorfos a él, si en la demostración del teorema se emplean solamente las propiedades de las operaciones y no las propiedades individuales de los elementos de este anillo. Por esta razón, *no vamos a considerar como diferentes los anillos o los campos que son isomorfos*; éstos serán para nosotros distintos ejemplares de un mismo anillo o campo.

Apliquemos este concepto al problema de la construcción del campo de los números complejos. La construcción del campo de los números complejos expuesta en el § 17, y basada en la aplicación de los puntos del plano, no es la única posible. En lugar de puntos se podrían haber tomado segmentos (vectores) en el plano que parten

del origen de coordenadas y, dando estos vectores por sus componentes a, b sobre los ejes coordenados, se determinaría la suma y el producto de vectores mediante las mismas fórmulas (2) y (3) del § 17, así como en el caso de los puntos del plano. En general, se podría no insistir en aplicar objetos geométricos. Observando que los puntos en el plano, así como los vectores en el plano, se determinan por pares ordenados de números reales (a, b), se puede tomar simplemente el conjunto de tales pares e introducir en él la suma y el producto según las fórmulas (2) y (3) del párrafo indicado.

En realidad, estos campos no se distinguirían por sus propiedades algebraicas, como muestra el teorema siguiente:

Todas las ampliaciones del campo de números reales D, obtenidas por adjunción al campo D de la raíz de la ecuación

$$x^2 + 1 = 0, \tag{1}$$

son isomorfas entre sí.

En efecto, sea dado algún campo P que represente una ampliación del campo D y que contenga al elemento que satisface a la ecuación (1). La elección de la notación de este elemento corre a nuestro cargo y, para este fin, emplearemos la letra i. Por lo tanto, se cumple la igualdad $i^2 + 1 = 0$ (de donde $i^2 = -1$), aquí la elevación a potencia y la suma se deben entender en el sentido de las operaciones definidas en el campo P. Queremos hallar ahora el campo $D(i)$ que se obtiene por adjunción del elemento i al campo D, es decir, hallar el subcampo mínimo del campo P que contiene al campo D y al elemento i.

Examinemos con este fin todos los elementos α del campo P que se pueden escribir de la forma

$$\alpha = a + bi, \tag{2}$$

donde a y b son números reales arbitrarios, y el producto del número b por el elemento i, así como la suma del número a y este producto, se deben entender en el sentido de las operaciones definidas en el campo P. Ningún elemento α del campo P puede poseer dos distintas expresiones de esta forma, puesto que de

$$\alpha = a + bi = \overline{a} + \overline{b}i,$$

siendo $b \neq \overline{b}$, resultaría

$$i = \frac{\overline{a} - a}{b - \overline{b}},$$

o sea, i sería un número real; si $b = \overline{b}$, resulta $a = \overline{a}$. En particular, entre los elementos del campo P que se expresan en la forma (2), figuran todos los números reales (cuando $b = 0$), y también el mismo elemento i (cuando $a = 0$, $b = 1$).

Demostremos que *el conjunto de todos los elementos de la forma* (2) *forman un subcampo del campo P*; éste será precisamente el campo buscado $D(i)$. Sean dados los elementos $\alpha = a + bi$ y $\beta = c + di$. Aplicando las leyes conmutativa y asociativa de la adición, así como la ley distributiva, que rigen en el campo P, obtenemos:

$$\alpha + \beta = (a + bi) + (c + di) = (a + c) + (bi + di),$$

de donde,

$$\alpha + \beta = (a + c) + (b + d)\, i, \tag{3}$$

o sea, esta suma pertenece de nuevo al conjunto de elementos considerado. Por otra parte,

$$-\beta = (-c) + (-d)\, i,$$

pues, en virtud de (3), se cumple la igualdad $\beta + (-\beta) = 0 + 0i = 0$; por lo tanto,

$$\alpha - \beta = \alpha + (-\beta) = (a - c) + (b - d)\, i, \tag{3'}$$

es decir, la resta no sale fuera de los límites del conjunto considerado. Aplicando de nuevo las propiedades I-V a que satisfacen las operaciones en el campo P (véase § 44) y basándose en la igualdad $i^2 = -1$, obtenemos:

$$\alpha\beta = (a + bi)(c + di) = ac + adi + bci + bdi^2,$$

o sea,

$$\alpha\beta = (ac - bd) + (ad + bc)\, i; \tag{4}$$

por lo tanto, el producto de dos elementos cualesquiera de la forma (2) es de nuevo un elemento de esta misma forma. Finalmente, supongamos que $\beta \neq 0$, es decir, que al menos uno de los números c, d, sea diferente de cero. Entonces también $c - di \neq 0$ y

$$(c + di)(c - di) = c^2 - (di)^2 = c^2 - d^2 i^2 = c^2 + d^2,$$

siendo $c^2 + d^2 \neq 0$. Por consiguiente, aplicando la afirmación hecha en el párrafo anterior, de que en cualquier campo se conservan todas las reglas de las operaciones con los quebrados y, por ende, un quebrado no varía al multiplicar su numerador y denominador por un elemento diferente de cero, obtenemos:

$$\frac{\alpha}{\beta} = \frac{a + bi}{c + di} = \frac{(a + bi)(c - di)}{(c + di)(c - di)} = \frac{(ac + bd) + (bc - ad)\, i}{c^2 + d^2},$$

es decir, el elemento

$$\frac{\alpha}{\beta} = \frac{ac + bd}{c^2 + d^2} + \frac{bc - ad}{c^2 + d^2}\, i \tag{4'}$$

tiene de nuevo la forma (2).

Demostremos ahora que *el subcampo obtenido $D(i)$ del campo P es isomorfo al campo de puntos del plano construido en el* § 17. Asocian-

do al elemento $a + bi$ del campo $D(i)$ el punto (a, b), en virtud de la unicidad de la expresión de la forma (2) para los elementos del campo $D(i)$, se obtiene una correspondencia biunívoca entre los elementos de este campo y todos los puntos del plano. En esta correspondencia, al número real a le corresponde el punto $(a, 0)$, debido a la igualdad $a = a + 0i$, y al elemento $i = 0 + 1 \cdot i$, el punto $(0, 1)$. Por otra parte, comparando las fórmulas (3) y (4) del presente párrafo con las fórmulas (2) y (3) del § 17, obtenemos que a la suma y al producto de los elementos α y β del campo $D(i)$ les corresponden los puntos que son la suma y, respectivamente, el producto de los puntos correspondientes de α y β.

Como todos los campos que son isomorfos a un campo dado son isomorfos entre sí el teorema queda demostrado. Vemos, en particular, que la elección de las fórmulas (2) y (3) en el § 17 para la definición de las operaciones con los puntos no fue casual y no puede ser modificada.

Además de los métodos de construcción del campo de los números complejos, examinados anteriormente, existen muchos otros métodos. Señalemos uno de estos, aplicando la suma y multiplicación de matrices.

Examinemos el anillo no conmutativo de las matrices de segundo orden sobre el campo de los números reales. Es evidente que las matrices escalares

$$\begin{pmatrix} a & 0 \\ 0 & a \end{pmatrix}$$

forman en este anillo un subcampo que es isomorfo al campo de los números reales. Pero, resulta que *en el anillo de las matrices de segundo orden sobre el campo de los números reales, se puede hallar también un subcampo que es isomorfo al campo de los números complejos.* En efecto, pongamos en correspondencia a cada número complejo $a + bi$ la *matriz*

$$\begin{pmatrix} a & b \\ -b & a \end{pmatrix}.$$

De este modo, resulta una aplicación biunívoca de todo el campo de números complejos en una parte del anillo de las matrices de segundo orden; además, de las igualdades

$$\begin{pmatrix} a & b \\ -b & a \end{pmatrix} + \begin{pmatrix} c & d \\ -d & c \end{pmatrix} = \begin{pmatrix} a+c & b+d \\ -(b+d) & a+c \end{pmatrix},$$

$$\begin{pmatrix} a & b \\ -b & a \end{pmatrix} \cdot \begin{pmatrix} c & d \\ -d & c \end{pmatrix} = \begin{pmatrix} ac-bd & ad+bc \\ -(ad+bc) & ac-bd \end{pmatrix}$$

se desprende que esta aplicación es isomorfa, puesto que las matrices que figuran en los segundos miembros de estas igualdades corresponden a los números complejos $(a + c) + (b + d)i = (a + bi) + (c + di)$ y $(ac - bd) + (ad + bc)i = (a + bi)(c + di)$. En particular la matriz

$$\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$$

desempeña el papel de unidad imaginaria.

El resultado obtenido señala otro posible método de construcción del campo de números complejos, que es tan satisfactorio como los considerados anteriormente.

§ 47. Algebra lineal y álgebra de los polinomios sobre un campo arbitrario

En los capítulos precedentes dedicados al **álgebra lineal**, el campo de números reales desempeñaba ordinariamente el papel de campo fundamental. No obstante, se comprueba sin dificultad alguna *que muchos teoremas de estos capítulos se generalizan palabra por* palabra al caso de un campo fundamental arbitrario.

Así, pues, *para un campo fundamental arbitrario P son válidos el método de Gauss de resolución de los sistemas de ecuaciones lineales, la teoría de los determinantes y la regla de Cramer*, expuestas en el cap. 1. Solamente la observación sobre los determinantes antisimétricos, expuesta al final del § 4, exige la suposición de que la característica del campo P **sea diferente de dos.** Por cierto, la demostración de la propiedad 4 de este mismo párrafo carece de valor si la característica del campo P es igual a dos, a pesar de que sea válida la propiedad misma.

Es conveniente señalar también que la afirmación, enunciada a menudo en el cap. 1, sobre la existencia de un conjunto infinito de soluciones distintas de un sistema indeterminado de ecuaciones lineales, es válida también en el caso de cualquier campo fundamental P **infinito**, pero carece de valor si el campo P es **finito**.

La teoría de la dependencia lineal de los vectores, la teoría del rango de una matriz y la teoría general de los sistemas de ecuaciones lineales, expuestas en el cap. 2, así como el álgebra de las matrices del cap. 3 se generalizan también totalmente al caso de un campo fundamental arbitrario.

La teoría general de las formas cuadráticas, expuesta en el § 26, *se generaliza al caso de cualquier campo fundamental P, cuya característica sea diferente de* **dos.** Sin esta restricción, pierde su valor el teorema fundamental de este párrafo.

Supongamos, por ejemplo, que $P = Z_2$, es decir, que es un campo constituido por dos elementos, 0 y 1, siendo $1 + 1 = 0$, de donde $-1 = 1$, y supongamos que sobre este campo se ha dado una forma cuadrática $f = x_1 x_2$. Si existe una transformación lineal

$$x_1 = b_{11}y_1 b_{12}y_2.$$

$$x_2 = b_{21}y_1 b_{22}y_2,$$

que lleve f a la forma canónica, en la igualdad

$$f = (b_{11}y_1 + b_{12}y_2)(b_{21}y_1 + b_{22}y_2) = b_{11}b_{21}y_1^2 + (b_{11}b_{22} + b_{12}b_{21})\, y_1 y_2 + b_{12}b_{22}y_2^2$$

tiene que ser igual a cero el coeficiente $b_{11}b_{22} + b_{12}b_{21}$ del producto $y_1 y_2$. Sin embargo, este coeficiente es igual al determinante de la transformación lineal considerada, pues, ya sea $b_{12}b_{21} = 1$, o $b_{12}b_{21} = 0$, en ambos casos, $b_{12}b_{21} = -b_{12}b_{21}$. Resulta que nuestra transformación lineal es degenerada.

El contenido ulterior del cap. 6 se refiere esencialmente a las formas cuadráticas de coeficientes complejos o reales.

Finalmente, *para el caso de un campo fundamental arbitrario P es válida toda la teoría de los espacios lineales y sus transformaciones lineales, expuesta en el cap. 7*. Por cierto, el concepto de raíz característica está ligado con la teoría de los polinomios sobre un campo arbitrario, de la que se hablará más adelante. Obsérvese que el teorema del § 33, sobre la relación entre las raíces características y los valores propios, se enuncia ahora del modo siguiente: las raíces características de una transformación lineal φ, pertenecientes al campo fundamental P, y sólo éstas, son valores propios de esta transformación.

La teoría de los espacios euclídeos (cap. 8) está ligada esencialmente con el campo de los números reales.

También se pueden generalizar para el caso de un campo fundamental arbitrario P algunos de los apartados del **álgebra de los polinomios** expuestos anteriormente. Sin embargo, es necesario fijar previamente el sentido exacto del concepto de polinomio sobre un campo arbitrario.

Esto se debe a que en el § 20 se señalaron dos puntos de vista sobre el concepto de polinomio: el concepto formal algebraico y el teórico funcional. Ambos se pueden generalizar al caso de un campo fundamental arbitrario. No obstante, siendo equivalentes para el caso de campos numéricos (véase el § 24) y, como fácilmente se comprueba para campos infinitos en general, *dejan de ser equivalentes ya para campos finitos*.

Veamos, por ejemplo, el campo Z_2 introducido en el § 45, compuesto de dos elementos 0 y 1, siendo $1+1=0$. Los polinomios $x+1$ y x^2+1 con coeficientes de este campo, son distintos, o sea, no satisfacen a la definición algebraica de igualdad de polinomios. Sin embargo, ambos polinomios toman el valor 1 para $x=0$ y el valor 0 para $x=1$, es decir, como «funciones» de la «variable» x, que toma valores en el campo Z_2, tienen que suponerse iguales.

En el campo Z_3 compuesto de tres elementos: 0, 1, 2, donde $1+2=0$, se encuentran en la misma situación los polinomios: x^3+x+1 y $2x+1$. En general, se pueden indicar ejemplos de este tipo para todos los campos finitos.

Por lo tanto, en la teoría relacionada al caso de un campo arbitrario P, es imposible admitir el punto de vista teórico funcional sobre los polinomios. Por consiguiente, es necesario aclarar definitivamente la definición formal algebraica de polinomio. Con este fin, realizaremos una **construcción del anillo de los polinomios sobre un campo arbitrario** P, que no utiliza desde el mismo comienzo la expresión ordinaria de los polinomios mediante «la indeterminada» x.

Examinemos todos los sistemas finitos ordenados posibles de elementos del campo P que tienen la forma

$$(a_0,\ a_1,\ \ldots,\ a_{n-1},\ a_n), \tag{1}$$

donde n es arbitrario, $n \geqslant 0$; para $n > 0$ tiene que ser $a_n \neq 0$. Determinando para los sistemas de la forma (1), la suma y el producto de acuerdo a las fórmulas (3) y (4) del § 20, convertimos el conjunto de estos sistemas en un anillo conmutativo; para demostrar que se cumplen las propiedades necesarias no hay más que repetir palabra por palabra lo que se hizo en el § 20 para los polinomios numéricos.

En el anillo que hemos construido, los sistemas de la forma (a) (caso de $n = 0$) forman un subcampo que es isomorfo al campo P. Esto permite identificar tales sistemas con los elementos correspondientes a del campo P, o sea, suponer

$$(a) = a \text{ para todos los } a \text{ de } P. \tag{2}$$

Por otra parte, designemos el sistema $(0,\ 1)$ con la letra x,

$$x = (0,\ 1).$$

Entonces, aplicando la definición de producto indicada anteriormente, obtenemos que $x^2 = (0,\ 0,\ 1)$ y, en general,

$$x^k = (\underbrace{0,\ 0,\ \ldots,\ 0}_{k \text{ veces}},\ 1) \tag{3}$$

Aplicando ahora la definición de suma y producto de sistemas *ordenados, y también las igualdades (2) y (3), resulta:*

$$(a_0,\ a_1,\ a_2,\ \ldots,\ a_{n-1},\ a_n) =$$

$$= (a_0) + (0,\ a_1) + (0,\ 0,\ a_2) + \ldots$$

$$\ldots + (\underbrace{0,\ 0,\ \ldots,\ 0}_{n-1 \text{ veces}},\ a_{n-1}) + (\underbrace{0,\ 0,\ \ldots,\ 0}_{n \text{ veces}},\ a_n) =$$

$$= (a_0) + (a_1)(0,\ 1) + (a_2)(0,\ 0,\ 1) + \ldots$$

$$\ldots + (a_{n-1})(\underbrace{0,\ 0,\ \ldots,\ 0}_{n-1 \text{ veces}},\ 1) + (a_n)(\underbrace{0,\ 0,\ \ldots,\ 0}_{n \text{ veces}},\ 1) =$$

$$= a_0 + a_1x + a_2x^2 + \ldots + a_{n-1}x^{n-1} + a_nx^n.$$

Por lo tanto, todo sistema ordenado de la forma (1) se puede expresar en forma de un polinomio con respecto a x, con coeficientes del campo P, siendo evidentemente esta expresión única. Basándose finalmente en la conmutatividad ya demostrada de la suma, se puede pasar a la expresión según las potencias decrecientes de x.

Por consiguiente, construimos aquí un anillo conmutativo que, naturalmente, se debe denominar *anillo de los polinomios en la indeterminada x sobre el campo P*. Este anillo se designa con la notación $P[x]$.

En el anillo $P[x]$ está contenido el mismo P, lo cual ya se había demostrado antes. Así como en el caso de anillos de polinomios sobre campos numéricos (véase § 20), *el anillo $P[x]$ posee unidad, no contiene divisores de cero y no es campo.*

Si el campo P está contenido en un campo más amplio $\overline{P}$, el anillo $P[x]$ es un subanillo del anillo $\overline{P}[x]$: puesto que todo polinomio con coeficientes de P se puede considerar como polinomio sobre el campo $\overline{P}$, y la suma y el producto de polinomios dependen sólo de sus coeficientes, no variando al pasar a un campo más amplio.

Para tener una idea mejor acerca del concepto verdadero del «anillo de los polinomios sobre el campo P», examinémoslo también desde otro ángulo.

Supongamos que el campo P está contenido como subanillo en algún anillo conmutativo L. Un elemento α del anillo L se llama *algebraico sobre el campo P*, si existe una ecuación de grado n, $n \geqslant 1$, con coeficientes del campo P, a la cual satisface el elemento α; si tal ecuación no existe, el elemento α se llama *trascendente sobre el campo P*. Está claro que el elemento x del anillo $P[x]$ es trascendente sobre el campo P.

Subsiste el **teorema** siguiente:

Si el elemento α del anillo L es trascendente sobre el campo P, el subanillo L', obtenido por adjunción del elemento α al campo P (o sea, el subanillo mínimo del anillo L que contiene al campo P y al elemento α), *es isomorfo al anillo de los polinomios $P[x]$.*

En efecto, cualquier elemento β del anillo L que se puede expresar en la forma

$$\beta = a_0\alpha^n + a_1\alpha^{n-1} + \ldots + a_{n-1}\alpha + a_n, \ n \geqslant 0, \tag{4}$$

con coeficientes $a_0, a_1, \ldots, a_{n-1}, a_n$ del campo P, estará contenido en el subanillo L'. El elemento β no puede poseer dos expresiones distintas de la forma (4), pues, restando una expresión de la otra resultaría una ecuación sobre el campo P a la que satisfaría el elemento α, lo cual contradice a la hipótesis de que el elemento α es trascendente. Sumando los elementos de la forma (4), según las reglas de la adición en el anillo L, se pueden sumar los coeficientes de potencias iguales de α; sin embargo, ésto coincide con la regla de adición de los polinomios. Por otra parte, multiplicando los elementos de la forma (4) según las leyes de la multiplicación en el anillo L y aplicando la ley de distribución, podemos efectuar la multiplicación término a término y reunir después los términos semejan-

tes; evidentemente, esto nos lleva a la conocida regla de la multiplicación de los polinomios. Con esto queda demostrado que los elementos de la forma (4) forman en el anillo L un subanillo que contiene al campo P y al elemento α, es decir, que coincide con L', y que este subanillo es isomorfo al anillo de los polinomios $P[x]$.

Vemos, pues, que la elección que hicimos de las definiciones para las operaciones con los polinomios no fue casual: ésta queda completamente determinada debido a que el elemento x del anillo $P[x]$ tiene que ser trascendente sobre el campo P.

Obsérvese que al construir el anillo de los polinomios $P[x]$ nunca se aplicó la división de los elementos del campo P y solamente una vez, cuando se demostraba la proposición sobre el grado del producto de los polinomios, hubo que referirse a la ausencia de divisores de cero *en el campo* P. Por consiguiente, *se puede tomar un anillo conmutativo arbitrario* L, y repitiendo la construcción realizada *anteriormente*, resulta *el anillo de los polinomios* $L[x]$ *sobre el anillo* L; **si en este caso el anillo L no contiene divisores de cero**, el grado *del producto de los polinomios será igual a la suma de los grados* de los factores y, por consiguiente, **el anillo de los polinomios $L[x]$ *tampoco contendrá divisores de cero.***

Volviendo a considerar los polinomios con coeficientes de un **campo** *arbitrario* P, *observemos* que, *substancialmente, toda la* **teoría de la divisibilidad de los polinomios** (véanse §§ 20-22) se generaliza a este caso. Precisamente, *en el anillo* $P[x]$ *tiene valor el algoritmo de la división con resto*, en la que el cociente, así como el residuo, pertenecen también al anillo $P[x]$. También tiene *sentido en el anillo* $P[x]$ *el concepto de divisor y se conservan todas sus propiedades principales*. Además, como el algoritmo de la división no nos saca fuera de los límites del campo fundamental P, se puede afirmar que *la propiedad del polinomio* $\varphi(x)$ *de ser divisor de* $f(x)$ *no depende de que se considere el campo* P *o cualquier ampliación de él.*

En el anillo $P[x]$ *se conservan también la definición y todas las propiedades del máximo común divisor, incluyendo el algoritmo de Euclides y el teorema demostrado en el § 21 mediante este algoritmo.* Obsérvese que, como el algoritmo de la división con resto no depende, como ya sabemos, del campo fundamental elegido, se puede afirmar que *el máximo común divisor de dos polinomios dados tampoco depende de que se considere el campo* P *o una ampliación arbitraria* $\overline{P}$ *del mismo*. Finalmente, *para los polinomios sobre el campo* P *tiene sentido el concepto de raíz y conservan su valor las propiedades fundamentales de las raíces.*

También se conserva la teoría de las raíces múltiples; por cierto, al final del párrafo siguiente volveremos a examinar esta cuestión.

Estas observaciones nos permitirán referirnos en adelante a los §§ 20-22 al estudiar los polinomios sobre cualquier campo P.

§ 48. Descomposición de los polinomios en factores irreducibles

En virtud del teorema de existencia de raíz (§ 24), para los campos de números complejos y reales quedó demostrada la existencia y unicidad de la descomposición de un polinomio en factores irreducibles. Estos resultados son casos particulares de los teoremas generales referentes a polinomios sobre un campo arbitrario P. El presente párrafo está dedicado a la exposición de esta teoría general, que es análoga a la teoría de la descomposición de los números enteros en factores primos.

Determinemos primero los polinomios que desempeñan en el anillo de los polinomios el mismo papel que los números primos en el anillo de los números enteros. Subrayemos previamente que en esta definición se va a tratar solamente de **polinomios de grado mayor o igual a la unidad**; esto corresponde al hecho de que en la definición de los números primos, al estudiar las descomposiciones de los números enteros en factores primos, los números 1 y -1 se excluyan.

Sea dado un polinomio $f(x)$ de grado n, $n \geqslant 1$, con coeficientes pertenecientes al campo P. En virtud de la propiedad V del § 21, todos los polinomios de grado cero son divisores de $f(x)$. Por otra parte, en virtud de VII, también son divisores de $f(x)$ todos los polinomios $cf(x)$, donde c es un elemento de P diferente de cero, agotándose con éstos todos los divisores de $f(x)$ de grado n. En cuanto a los divisores de $f(x)$, de grado mayor que 0, pero menor que n, éstos pueden existir en el anillo $P[x]$, o pueden no existir. En el primer caso, el polinomio $f(x)$ se llama *reducible* en el campo P (o sobre el campo P); en el segundo caso, *irreducible* en este campo (o sobre este campo).

Recordando la definición de divisor se puede decir que *un polinomio $f(x)$ de grado n es reducible en el campo P, si se puede descomponer sobre este campo (o sea, en el anillo $P[x]$) en el producto de dos factores de grados menores que n*:

$$f(x) = \varphi(x)\psi(x); \tag{1}$$

$f(x)$ es irreducible en el campo P, si en cualquiera de sus descomposiciones de la forma (1), *uno de los factores es de grado 0 y otro, de grado n.*

Es menester tener en cuenta que se puede hablar de reducibilidad o irreducibilidad de un polinomio solamente con respecto a un campo dado P, pues, un polinomio que es irreducible en este campo puede ser reducible en cierta ampliación $\overline{P}$ de él. Por ejemplo, el polinomio $x^2 - 2$ de coeficientes enteros es irreducible en el campo de números racionales, puesto que no se puede descomponer en un producto de dos factores de primer grado con coeficientes racionales.

Sin embargo, este polinomio es reducible en el campo de números reales, como muestra la igualdad

$$x^2-2=(x-\sqrt{2})(x+\sqrt{2}).$$

El polinomio x^2+1 no sólo es irreducible en el campo de números racionales, sino también en el campo de números reales; sin embargo, se hace reducible en el campo de números complejos, puesto que

$$x^2+1=(x-i)(x+i).$$

Indiquemos unas cuantas propiedades fundamentales de los polinomios irreducibles, recordando que se trata de polinomios irreducibles en el **campo** P.

α) *Todo polinomio de primer grado es irreducible.*

En efecto, si este polinomio se descompusiese en un producto de factores de menor grado, éstos tendrían que ser de grado cero. No obstante, el producto de cualesquiera polinomios de grado cero es de nuevo un polinomio de grado cero, y no de grado uno.

β) *Si el polinomio $p(x)$ es irreducible, lo es también cualquier polinomio $cp(x)$, donde c es un elemento de P diferente de cero.*

Esta propiedad es consecuencia de las propiedades I y VII § 21 y nos permitirá limitarnos, allí donde hiciese falta, al estudio de los polinomios irreducibles cuyos coeficientes superiores sean iguales a la unidad.

γ) *Si $f(x)$ es un polinomio arbitrario y $p(x)$ es un polinomio irreducible, entonces $f(x)$ es divisible por $p(x)$, o estos polinomios son primos entre sí.*

Si $(f(x), p(x))=d(x)$, el polinomio $d(x)$, siendo divisor del polinomio irreducible $p(x)$, es de grado 0 o bien es un polinomio de la forma $cp(x)$, $c\neq 0$. En el primer caso, $f(x)$ y $p(x)$ son primos entre sí, en el segundo, $f(x)$ es divisible por $p(x)$.

δ) *Si el producto de los polinomios $f(x)$ y $g(x)$ es divisible por un polinomio irreducible $p(x)$, al menos uno de estos factores es divisible por $p(x)$.*

En efecto, si $f(x)$ no es divisible por $p(x)$, según γ), $f(x)$ y $p(x)$ son primos entre sí, y, entonces, según la propiedad b) del § 21, el polinomio $g(x)$ tiene que ser divisible por $p(x)$.

La propiedad δ) se generaliza sin dificultad al caso del producto de cualquier número finito de factores.

Los dos teoremas que siguen son el objeto principal del presente párrafo.

Todo polinomio $f(x)$ de grado n, $n\geqslant 1$, del anillo $P[x]$, se descompone en un producto de factores irreducibles.

En efecto, si el mismo polinomio $f(x)$ es irreducible, el producto indicado consta de un solo factor. Si es reducible, se puede descomponer en un producto de factores de menor grado. Si entre estos factores

hay de nuevo reducibles, efectuamos la descomposición siguiente en factores, etc. Este proceso tiene que terminarse después de un número finito de ensayos, pues, sea cual fuese la descomposición de $f(x)$ en factores, la suma de sus grados tiene que ser igual a n, por lo que el número de factores que dependen de x no puede ser mayor que n.

La descomposición de los números enteros en factores primos es única, si nos limitamos a considerar los números enteros positivos. Sin embargo, en el anillo de todos los números enteros la unicidad subsiste, salvo el signo: así pues, $-6 = 2\cdot(-3) = (-2)\cdot 3$, $10 = 2\cdot 5 = (-2)(-5)$, etc. En el anillo de los polinomios nos encontramos con una situación análoga. Si

$$f(x) = p_1(x)\, p_2(x) \ldots p_s(x)$$

es una descomposición del polinomio $f(x)$ en un producto de factores irreducibles y si los elementos $c_1, c_2, \ldots, c_s$ del campo P son tales que su producto es igual a 1, entonces en virtud de β),

$$f(x) = [c_1p_1(x)]\cdot[c_2p_2(x)] \ldots [c_sp_s(x)]$$

también será una descomposición de $f(x)$ en un producto de factores irreducibles. Con éstas se agotan todas las descomposiciones de $f(x)$:

Si un polinomio $f(x)$ del anillo $P[x]$ se descompone de dos modos en un producto de factores irreducibles:

$$f(x) = p_1(x)\, p_2(x) \ldots p_s(x) = q_1(x)\, q_2(x) \ldots q_t(x), \tag{2}$$

entonces, $s = t$ y, con una numeración adecuada, se verifican las igualdades:

$$q_i(x) = c_ip_i(x), \quad i = 1, 2, \ldots, s, \tag{3}$$

donde c_i son elementos del campo P diferentes de cero.

Este teorema subsiste para los polinomios de primer grado, pues, éstos son irreducibles. Por lo tanto, la demostración se hará empleando el método de inducción sobre el grado del polinomio, es decir, se demostrará el teorema para $f(x)$, suponiendo que ya está demostrado para los polinomios de menor grado.

Como $q_1(x)$ es divisor de $f(x)$, en virtud de la propiedad δ) y de la igualdad (2), $q_1(x)$ será divisor por lo menos de uno de los polinomios $p_i(x)$, por ejemplo, de $p_1(x)$. Mas, como el polinomio $p_1(x)$ es irreducible y el grado de $p_1(x)$ es mayor que cero, existe un elemento c_1 tal que

$$q_1(x) = c_1p_1(x). \tag{4}$$

Poniendo en (2) esta expresión de $q_1(x)$ y simplificando por $p_1(x)$ (lo cual se permite, puesto que en el anillo $P[x]$ no hay divisores

de cero), se obtiene la igualdad

$$p_2(x)\, p_3(x) \ldots p_s(x) = [c_1 q_2(x)]\, q_3(x) \ldots q_t(x).$$

Como el grado del polinomio que es igual a estos productos, es menor que el grado de $f(x)$, queda ya demostrado que $s-1=t-1$, de donde $s=t$, y que existen unos elementos $c_2', c_3, \ldots, c_s$, tales que $c_2' p_2(x) = c_1 q_2(x)$, de donde $q_2(x) = (c_1^{-1} c_2')\, p_2(x)$, y $c_i p_i(x) = q_i(x)$, $i=3, \ldots, s$. Haciendo $c_1^{-1} c_2' = c_2$ y teniendo en cuenta (4), obtenemos la igualdad (3).

El teorema que acabamos de demostrar se puede enunciar más brevemente: *todo polinomio se descompone en factores irreducibles de un modo único, salvo factores de grado cero.*

Por cierto, siempre se puede considerar la descomposición de la siguiente forma especial, que *para cada polinomio ya es completamente única*: se toma cualquier descomposición del polinomio $f(x)$ en factores irreducibles y de cada uno de estos factores se saca fuera de paréntesis su coeficiente superior. Se obtiene la descomposición

$$f(x) = a_0 p_1(x)\, p_2(x) \ldots p_s(x), \tag{5}$$

donde todos los $p_i(x)$, $i=1, 2, \ldots, s$, son polinomios irreducibles cuyos coeficientes superiores son iguales a la unidad. El factor a_0 será igual al coeficiente superior del polinomio $f(x)$, lo que se *comprueba fácilmente efectuando las multiplicaciones en el segundo* miembro de la igualdad (5).

Los factores irreducibles que forman parte de la descomposición (5), no son todos necesariamente distintos. Si el polinomio irreducible $p(x)$ figura unas cuantas veces en la descomposición (5), se llama *factor múltiple* de $f(x)$: precisamente, k es el orden de multiplicidad, si en la descomposición (5) hay exactamente k factores iguales a $p(x)$. Si el factor $p(x)$ figura en (5) una sola vez, se llama *factor simple* (el orden de multiplicidad es igual a uno) de $f(x)$.

Si en la descomposición (5) los factores $p_1(x), p_2(x), \ldots, p_l(x)$ son distintos entre sí y cualquier otro factor es igual a uno de éstos, siendo k_i, $i=1, 2, \ldots, l$, el orden de multiplicidad del factor $p_i(x)$, la descomposición (5) se puede escribir en la forma siguiente:

$$f(x) = a_0 p_1^{k_1}(x)\, p_2^{k_2}(x) \ldots p_l^{k_l}(x). \tag{6}$$

A continuación se utilizará, por lo general, esta expresión, sin advertir ya que los exponentes son los órdenes de multiplicidad de los factores respectivos, es decir, que $p_i(x) \neq p_j(x)$ para $i \neq j$.

Dadas las descomposiciones de los polinomios $f(x)$ y $g(x)$ en factores irreducibles, el máximo común divisor $d(x)$ de estos polinomios es igual al producto de los factores que figuran simultáneamente en ambas descomposiciones, elevado cada factor a una potencia igual al mínimo de los órdenes de su multiplicidad en ambos polinomios.

En efecto, el producto indicado es divisor de cada uno de los polinomios $f(x)$ y $g(x)$, y, por esto, lo es también de $d(x)$. Si este producto fuese distinto de $d(x)$, en la descomposición de $d(x)$ en factores irreducibles estaría contenido un factor que no figura en la descomposición de alguno de los polinomios $f(x)$ y $g(x)$, lo cual es imposible, o bien, uno de los factores estaría elevado a una mayor potencia que la que tiene en la descomposición de alguno de los polinomios $f(x)$ y $g(x)$, lo cual, de nuevo, es imposible.

Este teorema es análogo a la regla según la cual se busca ordinariamente el máximo común divisor de los números enteros. Sin embargo, este teorema no puede sustituir al algoritmo de Euclides en el caso de los polinomios. En efecto, como sólo hay un número finito de números primos menores que un número entero positivo dado, la descomposición de un número entero en factores primos se consigue mediante un número finito de ensayos. Esto ya no se verifica en el anillo de los polinomios sobre un campo fundamental infinito y, en el caso general, no se puede señalar un método para la descomposición práctica de los polinomios en factores irreducibles. Incluso la resolución del problema para averiguar si el polinomio $f(x)$ es irreducible en un campo dado P, en el caso general, es muy difícil. Así pues, la descripción de todos los polinomios irreducibles para el caso de los campos de números complejos y reales fue obtenida en el § 24 como consecuencia de un teorema muy importante de existencia de la raíz. En lo que se refiere al campo de números racionales, se harán solamente algunas proposiciones de carácter particular en el § 56 con respecto a los polinomios irreducibles sobre este campo

Hemos demostrado que en el anillo de los polinomios, al igual que en el anillo de los números enteros, subsiste la descomposición en factores «primos» (irreducibles)* y que esta descomposición en cierto sentido es única. Surge la pregunta, ¿se pueden generalizar estos resultados a clases de anillos más amplios? En este caso nos limitaremos a considerar anillos conmutativos que posean unidad y no contengan divisores de cero.

Llamaremos *divisor de la unidad* a un elemento a del anillo para el que existe en este anillo el elemento recíproco a^{-1},

$$aa^{-1}=1.$$

En el anillo de los números enteros, éstos son los números 1 y -1; en el anillo de los polinomios $P[x]$, todos los polinomios de grado cero, o sea, los números del campo P, distintos de cero. Un elemento c diferente de cero, que no es divisor de la unidad, se llama elemento *primo* del anillo, si en cualquiera de sus descomposiciones en un producto de dos factores, $c=ab$, uno de estos factores es inevitablemente divisor de la unidad. En el anillo de los números enteros son elementos primos los números primos, en el anillo de los polinomios, los polinomios irreducibles.

¿Se descompone en un producto de factores primos cualquier elemento del anillo considerado, que no sea divisor de la unidad y sea diferente de cero? En

* Llamada descomposición factorial. (*Nota del T.*).

caso de afirmacion, ¿será única tal descomposición? Esto último hay que entenderlo en el sentido siguiente: si

$$a = p_1 p_2 \ldots p_k = q_1 q_2 \ldots q_l$$

son dos descomposiciones del elemento a en factores primos, entonces, $k = l$ y (posiblemente, después de cambiar la numeración)

$$q_i = p_i c_i, \quad i = 1, 2, \ldots, k,$$

donde c_i es divisor de la unidad.

En el caso general, ambas preguntas tienen una respuesta negativa. Aquí nos limitaremos con un ejemplo: **indicaremos un anillo en el que la descomposición en factores primos, aunque es posible, no es única.**

Examinemos los números complejos de la forma

$$\alpha = a + b\sqrt{-3}, \tag{7}$$

donde a y b son números enteros. Todos estos números forman un anillo sin divisores de cero que contiene la unidad; en efecto,

$$(a + b\sqrt{-3})(c + d\sqrt{-3}) = (ac - 3bd) + (bc + ad)\sqrt{-3}. \tag{8}$$

Llamemos *norma* del número $\alpha = a + b\sqrt{-3}$ al número entero positivo

$$N(\alpha) = a^2 + 3b^2.$$

En virtud de (8), *la norma del producto es igual al producto de las normas de los factores*:

$$N(\alpha\beta) = N(\alpha) N(\beta). \tag{9}$$

En efecto,

$$(ac - 3bd)^2 + 3(bc + ad)^2 = a^2c^2 + 9b^2c^2 + 3b^2c^2 + 3a^2d^2 =$$
$$= (a^2 + 3b^2)(c^2 + 3d^2).$$

Si en nuestro anillo el número α es divisor de la unidad, o sea, que el número α^{-1} también tiene la forma (7), entonces, por (9),

$$N(\alpha) \cdot N(\alpha^{-1}) = N(\alpha\alpha^{-1}) = N(1) = 1,$$

de aquí que, $N(\alpha) = 1$, pues los números $N(\alpha)$ y $N(\alpha^{-1})$ son enteros y positivos. Si $\alpha = a + b\sqrt{-3}$, de $N(\alpha) = 1$ se deduce que

$$N(\alpha) = a^2 + 3b^2 = 1;$$

sin embargo, esto es posible sólo cuando $b = 0$, $a = \pm 1$. Por lo tanto, *en nuestro anillo, así como en el anillo de los números enteros, son divisores de la unidad solamente los números* 1 *y* -1 *y solamente estos números tienen la norma igual a la unidad.*

Naturalmente, la igualdad (9) para la norma del producto se generaliza para el caso de un número finito de factores. De aquí fácilmente se deduce que *todo número* α *de nuestro anillo se puede descomponer en un producto de un número finito de factores primos*; la demostración se la dejamos al lector.

No obstante, *ya no se puede afirmar que la descomposición en factores primos es única.* Por ejemplo, se cumplen las igualdades.

$$4 = 2 \cdot 2 = (1 + \sqrt{-3})(1 - \sqrt{-3}).$$

En nuestro anillo no hay otros divisores de la unidad más que los números 1 y -1, por lo que el número $1 + \sqrt{-3}$ (así como el número $1 - \sqrt{-3}$) no puede diferenciarse del número 2 solamente en un factor que sea divisor de la unidad. No queda más que demostrar que en el anillo considerado, *cada uno de los números*

$2,\ 1+\sqrt{-3},\ 1-\sqrt{-3}$ *es primo.* En efecto, la norma de cada uno de estos tres números es igual al número 4. Sea α uno de estos números y supongamos que

$$\alpha=\beta\gamma.$$

Entonces, según (9), es posible uno de los tres casos:

1) $N(\beta)=4,\quad N(\gamma)=1$; 2) $N(\beta)=1,\quad N(\gamma)=4$;

3) $N(\beta)=N(\gamma)=2$.

En el primer caso, como ya sabemos, el número γ es divisor de la unidad, en el segundo caso, el divisor de la unidad es β. En lo que se refiere al tercer caso, éste es imposible, en general, puesto que, para enteros a y b, la igualdad

$$a^2+3b^2=2$$

es imposible.

Factores múltiples. A pesar de que más arriba se indicó que no sabemos descomponer los polinomios en factores irreducibles, existen métodos para saber si un polinomio dado posee factores múltiples o no, y, en caso de una respuesta positiva, éstos permiten reducir el estudio de este polinomio al estudio de otros que ya no poseen factores múltiples. Sin embargo, estos métodos exigen la imposición de ciertas restricciones al campo fundamental. Precisamente, todo el contenido ulterior del presente párrafo se va a exponer suponiendo que **la característica del campo P es cero.** Sin esta restricción, los teoremas sobre los factores múltiples que se demostrarán a continuación, pierden su valor; además, desde el punto de vista de las aplicaciones, el caso de campos de característica cero es el más importante, pues, incluye a todos los campos numéricos.

Obsérvese primero que el concepto de **derivada** de un polinomio, introducido en el § 22 para los polinomios de coeficientes complejos, y las propiedades principales de este concepto, también se generalizan para el caso considerado*. Demostremos ahora el siguiente teorema:

Si $p(x)$ es un factor irreducible múltiple de orden k, $k\geqslant 1$, del polinomio $f(x)$, entonces, es también un factor múltiple de orden $(k-1)$ de la derivada de este polinomio. En particular, un factor simple del polinomio no figura en la descomposición de la derivada.

En efecto, supongamos que

$$f(x)=p^k(x)\,g(x), \tag{10}$$

donde $g(x)$ ya no es divisible por $p(x)$. Derivando la igualdad (10), resulta:

$$f'(x)=p^k(x)\,g'(x)+kp^{k-1}(x)\,p'(x)\,g(x)=$$
$$=p^{k-1}(x)\,[p(x)\,g'(x)+kp'(x)\,g(x)].$$

El segundo término que figura entre paréntesis no es divisible por $p(x)$; en efecto, $g(x)$ no es divisible por $p(x)$ según la hipótesis y

* Para los campos de característica finita carece de valor la afirmación de que la derivada de un polinomio de grado n es de grado $n-1$.

$p'(x)$ es de grado menor, o sea, tampoco es divisible por $p(x)$. De aquí, en virtud de la irreducibilidad del polinomio $p(x)$ y de las *propiedades δ) del presente párrafo y IX del § 21, resulta nuestra* afirmación. Por otra parte, el primer término que figura entre corchetes es divisible por $p(x)$, por lo cual, toda esta suma no puede ser divisible por $p(x)$, o sea, el factor $p(x)$ figura, efectivamente, en $f'(x)$ con la multiplicidad $k-1$.

De nuestro teorema y del método indicado anteriormente para la averiguación del máximo común divisor de dos polinomios, se deduce que, dada la descomposición del polinomio $f(x)$ en factores irreducibles:

$$f(x)=a_0p_1^{k_1}(x)\,p_2^{k_2}(x)\,\ldots\,p_l^{k_l}(x), \tag{11}$$

el máximo común divisor del polinomio $f(x)$ y su derivada posee la siguiente descomposición en factores irreducibles:

$$(f(x),\ f'(x))=p_1^{k_1-1}(x)\,p_2^{k_2-1}(x)\,\ldots\,p_l^{k_l-1}(x), \tag{12}$$

*en la que, naturalmente, si $k_i=1$, el factor $p_i^{k_i-1}(x)$ se debe susti*tuir por la unidad. En particular, *el polinomio $f(x)$ no contiene factores múltiples cuando, y sólo cuando, éste es primo con su derivada.*

Por consiguiente, hemos aprendido a responder a la pregunta sobre la existencia de factores múltiples de un polinomio dado. Además, como la derivada de un polinomio, así como el máximo común divisor de dos polinomios, no dependen de que se considere *el campo P o cualquiera de sus ampliaciones $\overline{P}$, como consecuencia* del resultado que acabamos de demostrar obtenemos que:

Si un polinomio $f(x)$ con coeficientes de un campo P de característica cero, no tiene sobre este campo factores múltiples, no los tendrá tampoco sobre ninguna ampliación $\overline{P}$ del campo P.

En particular, si $f(x)$ es irreducible sobre P, y $\overline{P}$ es alguna ampliación del campo P, entonces, aunque $f(x)$ pueda ser ya reducible *sobre $\overline{P}$, no puede ser divisible por el cuadrado de un polinomio* irreducible (sobre $\overline{P}$).

Separación de los factores múltiples. Dado un polinomio $f(x)$ con la descomposición (11) y designando con $d_1(x)$ el máximo común divisor de $f(x)$ y su derivada $f'(x)$, la expresión (12) representa la descomposición de $d_1(x)$. Dividiendo (11) por (12), resulta:

$$v_1(x)=\frac{f(x)}{d_1(x)}=a_0p_1(x)\,p_2(x)\,\ldots\,p_l(x),$$

o sea, obtenemos un polinomio que carece de factores múltiples. Además, cualquier factor irreducible de $v_1(x)$ es también factor de $f(x)$. Con esto, la averiguación de los factores irreducibles de $f(x)$ se reduce a la averiguación de los mismos para el polinomio

$v_1(x)$, que, por lo general, es de menor grado y contiene solamente factores primos. Resolviendo este problema para $v_1(x)$, no queda más que determinar la multiplicidad en $f(x)$ de los factores irreducibles hallados, lo cual se consigue aplicando el algoritmo de la división.

Generalizando el método que acabamos de exponer, se puede pasar a examinar inmediatamente unos cuantos polinomios sin factores múltiples. Hallando sus factores irreducibles, no sólo obtenemos todos los factores irreducibles de $f(x)$, sino también sus órdenes de multiplicidad.

Sea (11) la descomposición de $f(x)$ en factores irreducibles y s, $s \geqslant 1$, el orden superior de multiplicidad de los factores. Designemos con $F_1(x)$ el producto de todos los factores de primer orden del polinomio $f(x)$; designemos con $F_2(x)$ el producto de todos los factores de segundo orden, pero tomados una sola vez cada uno, etc; finalmente, con $F_s(x)$ el producto de todos los factores de orden s, pero tomados también una sola vez cada uno. Si para cierto número j no existen en $f(x)$ factores de orden j, se supone que $F_j(x)=1$. Entonces, $f(x)$ es divisible por la k-ésima potencia del polinomio $F_k(x)$, $k=1, 2, \ldots, s$, y la descomposición (11) toma la forma

$$f(x)=a_0F_1(x)\,F_2^2(x)\,F_3^3(x)\ \ldots\ F_s^s(x),$$

y la descomposición (12) para $d_1(x)=(f(x),\ f'(x))$ se escribe así:

$$d_1(x)=F_2(x)\,F_3^2(x)\ \ldots\ F_s^{s-1}(x).$$

Designando con $d_2(x)$ el máximo común divisor del polinomio $d_1(x)$ y su derivada y, en general, con $d_k(x)$, el máximo común divisor de los polinomios $d_{k-1}(x)$ y $d'_{k-1}(x)$, obtenemos:

$$\begin{aligned} d_2(x)&=F_3(x)\,F_4^2(x)\ \ldots\ F_s^{s-2}(x),\\ d_3(x)&=F_4(x)\,F_5^2(x)\ \ldots\ F_s^{s-3}(x),\\ &\ldots\ldots\ldots\ldots\ldots\ldots\\ d_{s-1}(x)&=F_s(x),\\ d_s(x)&=1. \end{aligned}$$

De aquí que

$$\begin{aligned} v_1(x)&=\frac{f(x)}{d_1(x)}=a_0F_1(x)\,F_2(x)\,F_3(x)\ldots F_s(x),\\ v_2(x)&=\frac{d_1(x)}{d_2(x)}=F_2(x)\,F_3(x)\ldots F_s(x),\\ v_3(x)&=\frac{d_2(x)}{d_3(x)}=F_3(x)\ldots F_s(x),\\ &\ldots\ldots\ldots\ldots\ldots\ldots\\ v_s(x)&=\frac{d_{s-1}(x)}{d_s(x)}=F_s(x), \end{aligned}$$

y, finalmente,

$$F_1(x)=\frac{v_1(x)}{a_0v_2(x)},\quad F_2(x)=\frac{v_2(x)}{v_3(x)},\ \ldots,\ F_s(x)=v_s(x).$$

Por lo tanto, aplicando sólo métodos que no exigen el conocimiento de los factores irreducibles de $f(x)$, como la derivación, la aplicación del algoritmo de Euclides y del algoritmo de la división, se pueden hallar los polinomios

$F_1(x)$, $F_2(x)$, ..., $F_s(x)$ que carecen de factores múltiples, siendo cada factor irreducible del polinomio $F_k(x)$, $k = 1, 2, \ldots, s$, un factor de $f(x)$ de orden k.

El método expuesto no se puede considerar como método de descomposición de un polinomio en factores irreducibles, pues, para $s = 1$, es decir, para un polinomio sin factores múltiples se obtendría solamente $f(x) = F_1(x)$.

§ 49. Teorema de existencia de la raíz

El teorema fundamental demostrado en el § 23, sobre la existencia de una raíz en el campo de números complejos para cualquier polinomio numérico, no se puede generalizar para el caso de un campo arbitrario. En el presente párrafo se va a demostrar un teorema que, en cierta medida, sustituye en la teoría general de los campos al teorema fundamental indicado del álgebra de los números complejos.

Sea dado un polinomio $f(x)$ sobre el campo P. Surge la pregunta: ¿existe alguna ampliación $\overline{P}$ del campo P en la que $f(x)$ tenga ya por lo menos una raíz, si el polinomio $f(x)$ carece totalmente de raíces en el campo P? Aquí se puede suponer que el grado de $f(x)$ es mayor que la unidad, pues, para los polinomios de grado cero la pregunta no tiene sentido y cualquier polinomio de primer grado $ax + b$ tiene la raíz $-\frac{b}{a}$ en el mismo campo P. Por otra parte, está claro que podemos limitarnos al caso en que el polinomio $f(x)$ sea irreducible en el campo P, puesto que, en caso contrario, la raíz de cualquiera de sus factores irreducibles serviría de raíz para sí mismo.

La respuesta nos la da el siguiente **teorema de existencia de la raíz:**

Para cualquier polinomio $f(x)$, irreducible sobre el campo P, existe una ampliación de este campo en la que está contenida una raíz de $f(x)$. Todos los campos mínimos que contienen al campo P y a alguna raíz de este polinomio, son isomorfos entre sí.

Demostremos primero la segunda mitad del teorema.

Sea dado un polinomio

$$f(x) = a_0x^n + a_1x^{n-1} + \ldots + a_{n-1}x + a_n, \tag{1}$$

$n \geqslant 2$, *irreducible sobre* P, de modo que $f(x)$ *no tiene raíces en el* mismo campo P. Supongamos que existe una ampliación $\overline{P}$ del campo P, que contiene una raíz α de $f(x)$, y demostremos el siguiente **lema** que, además de ser necesario para nuestra demostración, es también de interés particular:

Si la raíz α, perteneciente a $\overline{P}$, de un polinomio $f(x)$ irreducible en P, es también raíz de un polinomio $g(x)$ del anillo $P[x]$, entonces $f(x)$ es un divisor de $g(x)$.

En efecto, en el campo $\overline{P}$ los polinomios $f(x)$ y $g(x)$ tienen un común divisor, $x - \alpha$, por lo que no son primos entre sí. No obstante

la propiedad de los polinomios de no ser primos entre sí no depende del campo que se haya elegido, por lo cual, se puede pasar al campo $\overline{P}$ y aplicar la propiedad γ) del párrafo anterior.

Hallemos ahora el subcampo mínimo $P(\alpha)$ del campo $\overline{P}$ que contiene al campo P y al elemento α. Ante todo, al campo buscado pertenecen todos los elementos de la forma

$$\beta = b_0 + b_1\alpha + b_2\alpha^2 + \ldots + b_{n-1}\alpha^{n-1}, \tag{2}$$

donde $b_0, b_1, b_2, \ldots, b_{n-1}$ son elementos del campo P. Ningún elemento del campo $\overline{P}$ puede poseer dos expresiones distintas de la forma (2), pues si se cumpliese también la igualdad

$$\beta = c_0 + c_1\alpha + c_2\alpha^2 + \ldots + c_{n-1}\alpha^{n-1},$$

donde por lo menos para un k fuese $c_k \neq b_k$, α sería raíz del polinomio

$$g(x) = (b_0 - c_0) + (b_1 - c_1)x + (b_2 - c_2)x^2 + \ldots + (b_{n-1} - c_{n-1})x^{n-1},$$

lo cual contradice al lema demostrado anteriormente, puesto que el grado de $g(x)$ es menor que el de $f(x)$.

Entre los elementos del campo $\overline{P}$ que tienen la forma (2), figuran todos los elementos del campo P (cuando $b_1 = b_2 = \ldots = b_{n-1} = 0$), y también el mismo elemento α (cuando $b_1 = 1$, $b_0 = b_2 = \ldots = b_{n-1} = 0$). Demostremos que *los elementos de la forma* (2) *forman todo el subcampo* $P(\alpha)$ *buscado*. En efecto, dados el elemento β (con la expresión (2)) y

$$\gamma = c_0 + c_1\alpha + c_2\alpha^2 + \ldots + c_{n-1}\alpha^{n-1},$$

en virtud de las propiedades de las operaciones en el campo $\overline{P}$, se tiene

$$\beta \pm \gamma = (b_0 \pm c_0) + (b_1 \pm c_1)\alpha + (b_2 \pm c_2)\alpha^2 + \ldots + (b_{n-1} \pm c_{n-1})\alpha^{n-1}$$

o sea, la suma y la diferencia de dos elementos cualesquiera de la forma (2) son de nuevo elementos de la misma forma.

Multiplicando β por γ resulta una expresión que contiene a α^n y a potencias más superiores de α. Sin embargo, de (1) y de la igualdad $f(\alpha) = 0$, se deduce que α^n y, por lo tanto, α^{n+1}, α^{n+2}, etc, se pueden expresar mediante potencias menores del elemento α. El método más simple para hallar la expresión de $\beta\gamma$ consiste en lo siguiente: sean

$$\varphi(x) = b_0 + b_1x + \ldots + b_{n-1}x^{n-1}, \quad \psi(x) = c_0 + c_1x + \ldots + c_{n-1}x^{n-1},$$

de donde, $\varphi(\alpha) = \beta$, $\psi(\alpha) = \gamma$. Multiplicando los polinomios $\varphi(x)$ y $\psi(x)$ y dividiendo este producto por $f(x)$, resulta

$$\varphi(x)\psi(x) = f(x)q(x) + r(x), \tag{3}$$

donde

$$r(x) = d_0 + d_1 x + \ldots + d_{n-1} x^{n-1}.$$

Hallando los valores de ambos miembros de la igualdad (3) para $x = \alpha$, resulta:

$$\varphi(\alpha)\,\psi(\alpha) = f(\alpha)\,q(\alpha) + r(\alpha),$$

o sea, como $f(\alpha) = 0$,

$$\beta\gamma = d_0 + d_1\alpha + \ldots + d_{n-1}\alpha^{n-1}.$$

Por lo tanto, el producto de dos elementos de la forma (2) es de nuevo un elemento de la misma forma.

Demostremos, finalmente, que si el elemento β *es de la forma* (2) y $\beta \neq 0$, el elemento β^{-1}, que existe en el campo $\overline{P}$, también se puede expresar en la forma (2). Para esto, tomemos el polinomio

$$\varphi(x) = b_0 + b_1 x + \ldots + b_{n-1} x^{n-1}$$

del anillo $P[x]$. Como el grado de $\varphi(x)$ es inferior al de $f(x)$ y el polinomio $f(x)$ es irreducible sobre P, los polinomios $\varphi(x)$ y $f(x)$ son primos entre sí y, en virtud de los §§ 21 y 47, en el anillo $P[x]$ existen unos polinomios $u(x)$ y $v(x)$ tales que

$$\varphi(x)\,u(x) + f(x)\,v(x) = 1;$$

además, se puede suponer que el grado de $u(x)$ es menor que n:

$$u(x) = s_0 + s_1 x + \ldots + s_{n-1} x^{n-1}.$$

De aquí, en virtud de la igualdad $f(\alpha) = 0$, resulta:

$$\varphi(\alpha)\,u(\alpha) = 1;$$

y, por esto, debido a la igualdad $\varphi(\alpha) = \beta$, se tiene:

$$\beta^{-1} = u(\alpha) = s_0 + s_1\alpha + \ldots + s_{n-1}\alpha^{n-1}.$$

Por lo tanto, el conjunto de los elementos del campo $\overline{P}$ que tienen la forma (2) forman un subcampo del campo $\overline{P}$; éste será el campo buscado $P(\alpha)$. Como hemos visto, para hallar la suma y el producto de los elementos β y γ de la forma (2) solamente hay que conocer los coeficientes de las expresiones de estos elementos mediante las potencias de α, por lo cual, se puede afirmar que subsiste el resultado siguiente: si además de $\overline{P}$ existe otra ampliación $\overline{P}'$ del campo P que contiene también una raíz α' del polinomio $f(x)$, y si $P(\alpha')$ es el subcampo mínimo del campo $\overline{P}'$ que contiene a P y a α', *los campos* $P(\alpha)$ *y* $P(\alpha')$ *son isomorfos*, donde, para obtener la correspondencia de isomorfismo entre ellos hay que asociar al elemento β de la forma (2) de $P(\alpha)$ el elemento

$$\beta' = b_0 + b_1\alpha' + b_2\alpha'^2 + \ldots + b_{n-1}\alpha'^{n-1}$$

de $P(\alpha')$ que tiene los mismos coeficientes. Con esto, queda demostrada la segunda mitad del teorema.

Pasemos ahora a demostrar la primera mitad de este teorema, que es la fundamental; lo expuesto anteriormente nos indicará el camino a seguir. Dado un polinomio $f(x)$ de grado $n \geqslant 2$, irreducible sobre el campo P, se necesita construir una ampliación del campo P que contenga una raíz de $f(x)$. Consideremos para esto todo el anillo de los polinomios $P[x]$ y dividámosle en clases disjuntas, incluyendo en una clase a los polinomios que al ser divididos por el polinomio dado $f(x)$ proporcionen residuos iguales. En otras palabras, los polinomios $\varphi(x)$ y $\psi(x)$ pertenecerán a una misma clase, si su diferencia es divisible por $f(x)$.

Convengamos en designar las clases obtenidas con las letras A, B, C, etc, y definamos la suma y el producto de las clases del siguiente modo. Tomemos dos clases cualesquiera A y B; elijamos en la clase A algún polinomio $\varphi_1(x)$, en la clase B, algún polinomio $\psi_1(x)$, y designemos con $\chi_1(x)$ la suma de estos polinomios,

$$\chi_1(x) = \varphi_1(x) + \psi_1(x),$$

y con $\theta_1(x)$, su producto

$$\theta_1(x) = \varphi_1(x) \cdot \psi_1(x).$$

Elijamos ahora en la clase A cualquier otro polinomio $\varphi_2(x)$, en la clase B, cualquier polinomio $\psi_2(x)$, y designemos respectivamente con $\chi_2(x)$ y $\theta_2(x)$ su suma y producto:

$$\chi_2(x) = \varphi_2(x) + \psi_2(x),$$
$$\theta_2(x) = \varphi_2(x) \cdot \psi_2(x).$$

Según la condición, los polinomios $\varphi_1(x)$ y $\varphi_2(x)$ pertenecen a una misma clase A, por lo cual su diferencia $\varphi_1(x) - \varphi_2(x)$ es divisible por $f(x)$; la diferencia $\psi_1(x) - \psi_2(x)$ posee esta misma propiedad. De aquí que la diferencia

$$\chi_1(x) - \chi_2(x) = [\varphi_1(x) + \psi_1(x)] - [\varphi_2(x) + \psi_2(x)] =$$
$$= [\varphi_1(x) - \varphi_2(x)] + [\psi_1(x) - \psi_2(x)] \qquad (4)$$

también es divisible por el polinomio $f(x)$. Esto mismo se cumple también para la diferencia $\theta_1(x) - \theta_2(x)$, puesto que

$$\theta_1(x) - \theta_2(x) = \varphi_1(x)\,\psi_1(x) - \varphi_2(x)\,\psi_2(x) =$$
$$= \varphi_1(x)\,\psi_1(x) - \varphi_1(x)\,\psi_2(x) + \varphi_1(x)\,\psi_2(x) - \varphi_2(x)\,\psi_2(x) =$$
$$= \varphi_1(x)\,[\psi_1(x) - \psi_2(x)] + [\varphi_1(x) - \varphi_2(x)]\,\psi_2(x). \qquad (5)$$

La igualdad (4) muestra que los polinomios $\chi_1(x)$ y $\chi_2(x)$ están situados en una misma clase. En otras palabras, la suma de cualquier

polinomio de la clase A y cualquier polinomio de la clase B pertenece a una clase C completamente determinada, que no depende de los *polinomios* que se hayan *elegido como «representantes»* de *las* clases A y B; llamemos a esta clase C, *suma* de las clases A y B:

$$C = A + B.$$

Análogamente, en virtud de (5), no depende tampoco de la elección de los representantes en las clases A y B la clase D, a la que pertenece el producto de cualquier polinomio de A por cualquier polinomio de B; llamemos a esta clase, *producto* de las clases A y B:

$$D = AB.$$

Demostremos que el conjunto de clases en que se ha dividido nuestro anillo de polinomios $P[x]$, después de haber introducido las operaciones anteriores de suma y producto, se convierte en un campo. En efecto, el cumplimiento de las **leyes asociativa** y **conmutativa** para ambas operaciones y de la **ley distributiva** es consecuencia de la subsistencia de estas leyes en el anillo $P[x]$, puesto que las operaciones con las clases se reducen a las operaciones con los polinomios situados en estas clases. Evidentemente, la clase formada por los polinomios que son divisibles por el polinomio $f(x)$ desempeña el papel del **cero**. Esta se denominará *clase cero* y se designará con el símbolo 0. La **opuesta** a la clase A, formada por los polinomios que al ser divididos por $f(x)$ dan un residuo $\varphi(x)$, será la clase formada por los polinomios que al ser divididos por $f(x)$ dan el residuo $-\varphi(x)$. De aquí se deduce que en el conjunto de los polinomios es posible la **resta**, siendo ésta unívoca.

Para demostrar que es posible la **división** en el conjunto de las *clases, hay que mostrar que existe una clase que desempeña el papel* de la unidad y que existe una clase recíproca para cualquier clase distinta de la clase cero. La **unidad** es evidentemente la clase de los polinomios que al ser divididos por $f(x)$ dan un residuo igual a 1; a ésta la llamaremos *clase unidad* y la designaremos con la notación E.

Sea dada ahora una clase A, distinta de la clase cero. Por consiguiente, un polinomio $\varphi(x)$, elegido en la clase A como representante, no será divisible por $f(x)$ y, como el polinomio $f(x)$ es irreducible, estos dos polinomios serán primos entre sí. Por lo tanto, en el anillo $P[x]$ existen unos polinomios $u(x)$ y $v(x)$ que satisfacen a la igualdad

$$\varphi(x)\,u(x) + f(x)\,v(x) = 1,$$

de donde

$$\varphi(x)\,u(x) = 1 - f(x)\,v(x). \tag{6}$$

El segundo miembro de la igualdad (6), al ser dividido por $f(x)$, da un residuo igual a 1 y, por lo tanto, pertenece a la clase unidad E.

Designando con B la clase a que pertenece el polinomio $u(x)$, la igualdad (6) muestra que

$$AB = E,$$

de donde $B = A^{-1}$. Con esto queda demostrada la existencia de una clase inversa para cualquier clase distinta de la clase cero, es decir, queda terminada la demostración de que las clases forman un campo.

Designemos este campo mediante $\overline{P}$ y demostremos que *éste es una ampliación del campo P*. A cada elemento a del campo P le corresponde la clase formada por los polinomios que al dividirlos por $f(x)$ dan un residuo igual a a; el mismo elemento a, considerado como un polinomio de grado cero, pertenece a esta clase. Todas las clases de esta forma especial forman en el campo $\overline{P}$ un subcampo isomorfo al campo P. En efecto, es evidente que la correspondencia es biunívoca; por otra parte, en estas clases se pueden elegir como representantes los elementos del campo P y, por consiguiente, a la suma (producto) de los elementos de P le va a corresponder la suma (producto) de las clases correspondientes. En consecuencia, tenemos derecho de no hacer diferencia entre los elementos del campo P y las clases que les corresponden.

Por último, designemos con X la clase formada por los polinomios que al ser divididos por $f(x)$ dan un residuo igual a x. Esta clase es un elemento del campo $\overline{P}$ completamente determinado, y queremos demostrar que *este elemento es raíz del polinomio $f(x)$*. Sea

$$f(x) = a_0x^n + a_1x^{n-1} + \ldots + a_{n-1}x + a_n.$$

Designemos mediante A_i la clase que corresponde, en el sentido indicado anteriormente, al elemento a_i del campo P, $i = 0, 1, \ldots$ $\ldots, n$ y veamos a qué es igual el elemento

$$A_0X^n + A_1X^{n-1} + \ldots + A_{n-1}X + A_n \tag{7}$$

del campo $\overline{P}$. Tomando los elementos a_i, $i = 0, 1, \ldots, n$, por representantes de las clases A_i y el polinomio x por representante de la clase X, y aplicando la definición de suma y producto de las clases, obtenemos que el mismo polinomio $f(x)$ está contenido en la clase (7). Pero, $f(x)$ es divisible por sí mismo, de donde resulta que (7) es la clase cero. Sustituyendo en (7) las clases A_i por sus elementos correspondientes a_i del campo P, obtenemos que en el campo $\overline{P}$ se verifica la igualdad

$$a_0X^n + a_1X^{n-1} + \ldots + a_{n-1}X + a_n = 0,$$

o sea, la clase X es verdaderamente raíz del polinomio $f(x)$.

Con esto queda terminada la demostración del teorema de existencia de la raíz. Obsérvese que, **tomando por P el campo de los números**

reales y poniendo $f(x) = x^2 + 1$, resulta otro método más de construcción del campo de los números complejos.

Del teorema de existencia de la raíz se pueden deducir consecuencias análogas a las que se dedujeron del teorema fundamental del álgebra de los números complejos (véase § 24). Hagamos primero una observación. Como cada factor lineal $x - c$ del polinomio $f(x)$ es irreducible, éste tiene que figurar en la descomposición **única** en factores irreducibles que posee $f(x)$.

Sin embargo, el número de factores lineales que hay en la descomposición de $f(x)$ en factores irreducibles no puede superar al grado de este polinomio. Así, llegamos al siguiente resultado:

Un polinomio $f(x)$ de grado n no puede tener en el campo P más de n raíces, incluso si cada raíz se cuenta tantas veces como indique su orden de multiplicidad.

Llamemos *campo de descomposición* del polinomio $f(x)$ de grado n sobre el campo P a una ampliación Q del campo P en la que estén contenidas n raíces de $f(x)$ (contando las raíces múltiples tantas veces cuantas indiquen sus órdenes de multiplicidad). Por consiguiente, el polinomio $f(x)$ se descompone sobre el campo Q en factores lineales. Además, ninguna otra ampliación del campo Q puede dar lugar a la aparición de nuevas raíces de $f(x)$.

Para cualquier polinomio $f(x)$ del anillo $P[x]$, existe sobre el campo P un campo de descomposición.

En efecto, si el polinomio $f(x)$ de grado n, $n \geqslant 1$, tiene n raíces en el mismo campo P, éste será el campo buscado de descomposición. Si $f(x)$ no se descompone sobre P en factores lineales, tomamos uno de sus factores irreducibles no lineales $\varphi(x)$ y, basándose en el teorema de la existencia de la raíz, ampliamos P hasta obtener un campo P' que contenga una raíz de $\varphi(x)$. Si el polinomio $f(x)$ no se descompone todavía sobre P' en factores lineales, ampliamos de nuevo el campo, creando una raíz para otro de los factores irreducibles no lineales que queden. Evidentemente, después de un número finito de operaciones, llegaremos a obtener para $f(x)$ un campo de descomposición.

Está claro que $f(x)$ puede poseer muchos campos distintos de descomposición. Se podría demostrar que todos los campos mínimos que contienen al campo P y a las n raíces del polinomio $f(x)$ (donde n es el grado del polinomio), son isomorfos entre sí. Como esta proposición no va a ser utilizada a continuación, no expondremos su demostración.

Raíces múltiples. En el párrafo anterior se había demostrado que un polinomio $f(x)$, dado sobre un campo P de característica 0, no tiene factores múltiples cuando, y sólo cuando, es primo con su derivada; también se había señalado que la carencia de factores múltiples de $f(x)$ sobre el campo P implica la carencia de factores

de este tipo sobre cualquier ampliación $\overline{P}$ del campo P. Aplicando esto al caso en que $\overline{P}$ sea un campo de descomposición para $f(x)$ y recordando la definición de raíz múltiple, llegamos al resultado siguiente:

Si un polinomio $f(x)$, dado sobre un campo P de característica 0, no tiene raíces múltiples en un campo dado de descomposición, éste es primo con su derivada $f'(x)$. Recíprocamente, si $f(x)$ es primo con su derivada, entonces no tiene raíces múltiples en ninguno de sus campos de descomposición.

En particular, de aquí se deduce que *un polinomio $f(x)$ irreducible sobre un campo P de característica 0, no puede tener raíces múltiples en ninguna ampliación de este campo.* Esta proposición deja de ser cierta para los campos de característica finita, circunstancia que desempeña un papel notable en la teoría general de los campos.

En conclusión, obsérvese que *para el caso de un campo arbitrario se conservan también las fórmulas de Vieta* (véase el § 24); en este caso, las raíces del polinomio se toman en un campo de descomposición del mismo.

§ 50. Campo de fracciones racionales

La teoría de las fracciones racionales, expuesta en el § 25, *se conserva también totalmente en el caso de un campo fundamental arbitrario.* Mas al pasar del campo de los números reales a un campo arbitrario P, el punto de vista según el cual la expresión $\frac{f(x)}{g(x)}$ se considera como una **función** de la variable x, tiene que ser desechado, puesto que, como ya se sabe, éste ya no es aplicable a los polinomios. Ante nosotros se plantea el problema de determinar el sentido que hay que atribuir a estas expresiones cuando los coeficientes pertenecen a un campo arbitrario P. Más exactamente, queremos construir un campo en el que esté contenido el anillo de los polinomios $P[x]$, pero de modo que las operaciones de suma y producto definidas en este nuevo campo, al ser aplicadas a los polinomios, coincidan con las operaciones en el anillo $P[x]$; abreviando, el anillo $P[x]$ tiene que ser un **subanillo** de este campo nuevo. Por otra parte, todo elemento de este nuevo campo tiene que representarse (en el sentido de la división en este campo) en forma de un cociente de dos polinomios. Como ahora se enseñará, tal campo puede ser construido para cualquier P; se designará con la notación $P(x)$ (obsérvese que la indeterminada está encerrada entre paréntesis) y se llamará *campo de fracciones racionales* sobre el campo P.

Supongamos primero que el anillo $P[x]$ ya es un subanillo de cierto campo Q. Si $f(x)$ y $g(x)$ son polinomios arbitrarios de $P[x]$, siendo $g(x) \neq 0$, existe en el campo Q un elemento unívocamente determinado, igual al cociente de la división de $f(x)$ por $g(x)$.

Designando este elemento, como ordinariamente se hace en el caso de un campo, mediante $\frac{f(x)}{g(x)}$, en virtud de la definición del cociente, se puede escribir la igualdad

$$f(x) = g(x) \cdot \frac{f(x)}{g(x)}, \tag{1}$$

donde el producto se debe entender en el sentido de la multiplicación en el campo Q. Puede ocurrir que algunos cocientes $\frac{f(x)}{g(x)}$ y $\frac{\varphi(x)}{\psi(x)}$ sean unos mismos elementos del campo Q; la condición para esto es la condición ordinaria de igualdad de las fracciones:

$$\frac{f(x)}{g(x)} = \frac{\varphi(x)}{\psi(x)} \textit{ cuando, y sólo cuando, } f(x)\,\psi(x) = \varphi(x)\,g(x).$$

En efecto, si $\frac{f(x)}{g(x)} = \frac{\varphi(x)}{\psi(x)} = \alpha$, en virtud de (1),

$$f(x) = g(x)\,\alpha, \quad \varphi(x) = \psi(x)\,\alpha,$$

de donde

$$f(x)\,\psi(x) = g(x)\,\psi(x)\,\alpha = g(x)\,\varphi(x).$$

Recíprocamente, si $f(x)\,\psi(x) = g(x)\,\varphi(x) = u(x)$ en el sentido de la multiplicación en el anillo $P[x]$, pasando al campo Q obtenemos las igualdades

$$\frac{f(x)}{g(x)} = \frac{u(x)}{g(x)\,\psi(x)} = \frac{\varphi(x)}{\psi(x)}.$$

Fácilmente se ve luego que la suma y el producto de cualesquiera elementos de Q, que son cocientes de polinomios de $P[x]$, se pueden representar de nuevo en forma de tales cocientes, cumpliéndose además las reglas comunes de adición y multiplicación de fracciones:

$$\frac{f(x)}{g(x)} + \frac{\varphi(x)}{\psi(x)} = \frac{f(x)\,\psi(x) + g(x)\,\varphi(x)}{g(x)\,\psi(x)}, \tag{2}$$

$$\frac{f(x)}{g(x)} \cdot \frac{\varphi(x)}{\psi(x)} = \frac{f(x)\cdot\varphi(x)}{g(x)\cdot\psi(x)}. \tag{3}$$

En efecto, multiplicando ambos miembros de cada una de estas igualdades por el producto $g(x)\,\psi(x)$ y aplicando (1), obtenemos igualdades válidas en el anillo $P[x]$. La subsistencia de las igualdades (2) y (3) se deduce ahora de que, debido a la ausencia de divisores de cero en el campo Q, ambos miembros de cada una de las igualdades obtenidas se pueden simplificar por el elemento $g(x)\,\psi(x)$, diferente de cero, sin infringir las igualdades.

Estas observaciones previas señalan el camino que hay que seguir para construir el campo $P(x)$. Sea dado un campo arbitrario P y sobre él, el anillo de los polinomios $P[x]$. A cada par ordenado

de polinomios $f(x)$, $g(x)$, donde $g(x) \neq 0$, ponemos en correspondencia el símbolo $\frac{f(x)}{g(x)}$, denominado *fracción racional con numerador* $f(x)$ *y denominador* $g(x)$. Subrayemos que esto es, simplemente, un símbolo que corresponde al par dado de polinomios, pues, por lo general, la división de los polinomios en el anillo mismo $P[x]$ no es posible y, por ahora, el anillo $P[x]$ no está contenido en ningún campo; incluso si $g(x)$ fuese divisor de $f(x)$, el nuevo símbolo $\frac{f(x)}{g(x)}$ se debe distinguir por ahora del polinomio que se obtiene como cociente al dividir $f(x)$ por $g(x)$.

Llamemos ahora *iguales* a las fracciones racionales $\frac{f(x)}{g(x)}$ y $\frac{\varphi(x)}{\psi(x)}$:

$$\frac{f(x)}{g(x)} = \frac{\varphi(x)}{\psi(x)}, \tag{4}$$

si en el anillo $P[x]$ se cumple la igualdad $f(x)\psi(x) = g(x)\varphi(x)$. Es evidente que cualquier fracción es igual a sí misma, y también que si una fracción es igual a otra, la segunda es igual a la primera. Demostremos que para este concepto de igualdad se cumple la **ley transitiva**. Sean dadas las igualdades (4) y

$$\frac{\varphi(x)}{\psi(x)} = \frac{u(x)}{v(x)}. \tag{5}$$

De las igualdades equivalentes a éstas en el anillo $P[x]$

$$f(x)\psi(x) = g(x)\varphi(x), \quad \varphi(x)v(x) = \psi(x)u(x)$$

se deduce que

$$f(x)v(x)\psi(x) = g(x)\varphi(x)v(x) = g(x)u(x)\psi(x);$$

por consiguiente, después de simplificar por el polinomio $\psi(x)$, distinto de cero (como denominador de una de las fracciones), resulta:

$$f(x)v(x) = g(x)u(x),$$

de donde, por la definición de igualdad de las fracciones,

$$\frac{f(x)}{g(x)} = \frac{u(x)}{v(x)},$$

que es lo que se quería demostrar.

Reunamos ahora en una clase todas las fracciones que sean iguales a una dada, las cuales, en virtud de la ley transitiva de la igualdad, serán iguales entre sí. Si en una clase hay por lo menos una fracción que no está contenida en otra clase, entonces, como se deduce de la ley transitiva de la igualdad, estas dos clases no tienen ningún elemento común.

Por lo tanto, el conjunto de todas las fracciones racionales, *escritas mediante los polinomios del anillo* $P[x]$, *se descompone en* clases disjuntas de fracciones iguales entre sí. Ahora queremos definir las operaciones algebraicas en este **conjunto de clases** de fracciones iguales, de modo que éste sea un campo. Para esto, vamos a definir las operaciones con las fracciones racionales y vamos a comprobar cada vez que la sustitución de los términos (o de los factores) por fracciones iguales a los mismos sustituye también la suma (o el producto) por una fracción igual. Esto permitirá hablar de la suma y producto de clases de fracciones iguales.

Hagamos previamente la siguiente observación que se aplicará a continuación: *una fracción racional se convierte en una fracción igual, si su numerador y denominador se multiplican por un mismo polinomio, diferente de cero, o si se simplifican por cualquier factor común.* En efecto,

$$\frac{f(x)}{g(x)} = \frac{f(x)\,h(x)}{g(x)\,h(x)},$$

pues en el anillo $P[x]$,

$$f(x)\,[g(x)\,h(x)] = g(x)\,[f(x)\,h(x)].$$

Definamos la *suma* de fracciones racionales por la fórmula (2); como $g(x) \neq 0$ y $\psi(x) \neq 0$, resulta que $g(x)\,\psi(x) \neq 0$, y el segundo miembro de esta fórmula es, evidentemente, una fracción racional. Si se ha dado que

$$\frac{f(x)}{g(x)} = \frac{f_0(x)}{g_0(x)}, \quad \frac{\varphi(x)}{\psi(x)} = \frac{\varphi_0(x)}{\psi_0(x)},$$

o sea, que

$$f(x)\,g_0(x) = g(x)\,f_0(x), \quad \varphi(x)\,\psi_0(x) = \psi(x)\,\varphi_0(x), \tag{6}$$

entonces, multiplicando ambos miembros de la primera de las igualdades (6) por $\psi(x)\,\psi_0(x)$, ambos miembros de la segunda por $g(x)\,g_0(x)$ y sumando después término a término estas igualdades, obtenemos:

$$[f(x)\,\psi(x) + g(x)\,\varphi(x)]\,g_0(x)\,\psi_0(x) =$$
$$= [f_0(x)\,\psi_0(x) + g_0(x)\,\varphi_0(x)]\,g(x)\,\psi(x),$$

lo cual es equivalente a la igualdad

$$\frac{f(x)\,\psi(x) + g(x)\,\varphi(x)}{g(x)\,\psi(x)} = \frac{f_0(x)\,\psi_0(x) + g_0(x)\,\varphi_0(x)}{g_0(x)\,\psi_0(x)}.$$

Por lo tanto, dadas dos clases de fracciones iguales entre sí, las sumas de cualquier fracción de una clase y cualquier fracción de otra clase son todas iguales entre sí, es decir, están situadas en una tercera clase completamente determinada. Esta clase se llama *suma* de las dos clases dadas.

La **conmutatividad** de esta suma es consecuencia inmediata de (2); la **asociatividad** se demuestra del modo siguiente:

$$\left[\frac{f(x)}{g(x)}+\frac{\varphi(x)}{\psi(x)}\right]+\frac{u(x)}{v(x)}=\frac{f(x)\psi(x)+g(x)\varphi(x)}{g(x)\psi(x)}+\frac{u(x)}{v(x)}=$$
$$=\frac{f(x)\psi(x)v(x)+g(x)\varphi(x)v(x)+g(x)\psi(x)u(x)}{g(x)\psi(x)v(x)}=$$
$$=\frac{f(x)}{g(x)}+\frac{\varphi(x)v(x)+\psi(x)u(x)}{\psi(x)v(x)}=\frac{f(x)}{g(x)}+\left[\frac{\varphi(x)}{\psi(x)}+\frac{u(x)}{v(x)}\right].$$

De la definición de igualdad de fracciones se deduce fácilmente que todas las fracciones de la forma $\frac{0}{g(x)}$, o sea, las fracciones con el numerador igual a cero, son iguales entre sí y forman una clase completa de fracciones iguales. A ésta la llamaremos *clase cero*; demostremos que esta clase desempeña en nuestra suma el papel del cero. En efecto, dada una fracción arbitraria $\frac{\varphi(x)}{\psi(x)}$, se tiene

$$\frac{0}{g(x)}+\frac{\varphi(x)}{\psi(x)}=\frac{0\cdot\psi(x)+g(x)\varphi(x)}{g(x)\psi(x)}=\frac{g(x)\varphi(x)}{g(x)\psi(x)}=\frac{\varphi(x)}{\psi(x)}.$$

De la igualdad

$$\frac{f(x)}{g(x)}+\frac{-f(x)}{g(x)}=\frac{0}{g^2(x)},$$

cuyo segundo miembro pertenece a la clase cero, se deduce ahora que la clase de las fracciones, iguales a la fracción $\frac{-f(x)}{g(x)}$, es la *opuesta* a la clase de las fracciones que son iguales a la fracción $\frac{f(x)}{g(x)}$. Como ya sabemos, de aquí se deduce la posibilidad de la *resta* unívoca.

Determinemos el *producto* de fracciones racionales por la fórmula (3). Como $g(x)\psi(x)\neq 0$, el segundo miembro de esta fórmula es, evidentemente, una fracción racional. Si, luego,

$$\frac{f(x)}{g(x)}=\frac{f_0(x)}{g_0(x)},\quad \frac{\varphi(x)}{\psi(x)}=\frac{\varphi_0(x)}{\psi_0(x)},$$

o sea,

$$f(x)g_0(x)=g(x)f_0(x),\quad \varphi(x)\psi_0(x)=\psi(x)\varphi_0(x),$$

entonces, multiplicando término a término estas últimas igualdades, obtenemos:

$$f(x)g_0(x)\varphi(x)\psi_0(x)=g(x)f_0(x)\psi(x)\varphi_0(x),$$

lo cual es equivalente a la igualdad

$$\frac{f(x)\varphi(x)}{g(x)\psi(x)}=\frac{f_0(x)\varphi_0(x)}{g_0(x)\psi_0(x)}.$$

Por lo tanto, por analogía con la definición de la suma de las clases, dada anteriomente, se puede hablar del *producto* de clases de fracciones iguales entre sí.

La **conmutatividad** y **asociatividad** de este producto es consecuencia directa de (3). El cumplimiento de la **ley distributiva** se demuestra del modo siguiente:

$$\left[\frac{f(x)}{g(x)}+\frac{\varphi(x)}{\psi(x)}\right]\frac{u(x)}{v(x)}=\frac{f(x)\psi(x)+g(x)\varphi(x)}{g(x)\psi(x)}\cdot\frac{u(x)}{v(x)}=$$

$$=\frac{[f(x)\psi(x)+g(x)\varphi(x)]\,u(x)}{g(x)\psi(x)v(x)}=\frac{f(x)\psi(x)u(x)+g(x)\varphi(x)u(x)}{g(x)\psi(x)v(x)}=$$

$$=\frac{f(x)\psi(x)u(x)v(x)+g(x)\varphi(x)u(x)v(x)}{g(x)\psi(x)v^2(x)}=\frac{f(x)u(x)}{g(x)v(x)}+\frac{\varphi(x)u(x)}{\psi(x)v(x)}=$$

$$=\frac{f(x)}{g(x)}\cdot\frac{u(x)}{v(x)}+\frac{\varphi(x)}{\psi(x)}\cdot\frac{u(x)}{v(x)}.$$

Fácilmente se observa que las fracciones de la forma $\frac{f(x)}{f(x)}$, o sea, las fracciones cuyos numeradores son iguales a sus denominadores, son iguales entre sí y forman una clase individual. Esta se llama *clase unidad* y en nuestra multiplicación desempeña el papel de la unidad:

$$\frac{f(x)}{f(x)}\cdot\frac{\varphi(x)}{\psi(x)}=\frac{f(x)\varphi(x)}{f(x)\psi(x)}=\frac{\varphi(x)}{\psi(x)}.$$

Si, finalmente, la fracción $\frac{f(x)}{g(x)}$ no pertenece a la clase cero, o sea, $f(x)\neq 0$, existe la **fracción** $\frac{g(x)}{f(x)}$. Como

$$\frac{f(x)}{g(x)}\cdot\frac{g(x)}{f(x)}=\frac{f(x)g(x)}{g(x)f(x)},$$

y el segundo miembro de esta igualdad pertenece a la clase unidad, la clase de las fracciones, iguales a la fracción $\frac{g(x)}{f(x)}$, será *repíproca* a la clase de las fracciones, iguales a la fracción $\frac{f(x)}{g(x)}$. De aquí se deduce que es posible la *división* unívoca.

Por lo tanto, *en virtud de las definiciones anteriores de las operaciones, las clases de fracciones racionales, iguales entre sí, con coeficientes del campo P, forman un campo conmutativo.* Este es el campo buscado $P(x)$. Por cierto, todavía tenemos que demostrar que en el campo construido está contenido un subanillo, isomorfo al anillo $P[x]$, y que cada elemento del campo se representa en forma de un cociente de dos elementos de este subanillo.

Si a un polinomio arbitrario $f(x)$ del anillo $P[x]$ ponemos en correspondencia la clase de fracciones racionales, iguales a la frac-

ción $\frac{f(x)}{1}$ (naturalmente, entre éstas también están contenidas las fracciones cuyos denominadores son iguales a la unidad), obtenemos una *aplicación biyectiva* del anillo $P[x]$ en el interior del campo que hemos construido. En efecto, de la igualdad

$$\frac{f(x)}{1}=\frac{\varphi(x)}{1}$$

resultaría $f(x)\cdot 1=1\cdot\varphi(x)$, o sea, $f(x)=\varphi(x)$. Como muestran las igualdades

$$\frac{f(x)}{1}+\frac{g(x)}{1}=\frac{f(x)\cdot 1+g(x)\cdot 1}{1^2}=\frac{f(x)+g(x)}{1},$$

$$\frac{f(x)}{1}\cdot\frac{g(x)}{1}=\frac{f(x)\cdot g(x)}{1},$$

esta aplicación es incluso un isomorfismo.

Por lo tanto, *las clases de fracciones, iguales a las fracciones de la forma* $\frac{f(x)}{1}$, *forman en nuestro campo un subanillo que es isomorfo al anillo* $P[x]$. Por esto, la fracción $\frac{f(x)}{1}$ se puede designar simplemente mediante $f(x)$. Finalmente, como la clase de las fracciones, iguales a la fracción $\frac{1}{g(x)}$, siendo $g(x)\neq 0$, es recíproca a la clase de las fracciones, iguales a la fracción $\frac{g(x)}{1}$, de la igualdad

$$\frac{f(x)}{1}\cdot\frac{1}{g(x)}=\frac{f(x)}{g(x)}$$

se deduce que *todos los elementos de nuestro campo se pueden considerar* (en el sentido de las operaciones definidas en este campo) *como cocientes de polinomios del anillo* $P[x]$.

De este modo, hemos construido el campo de fracciones racionales $P(x)$ sobre un campo arbitrario P. Tomando el anillo de los números enteros, en lugar del anillo de los polinomios, se puede construir de este mismo modo el campo de los números racionales. Agrupando estos dos casos y aplicando un método igual, se podría demostrar el teorema de que, en general, cualquier anillo conmutativo sin divisores de cero es un subanillo de algún campo.

CAPITULO XI

POLINOMIOS EN VARIAS INDETERMINADAS

§ 51. Anillo de los polinomios en varias indeterminadas

A veces se suelen considerar polinomios que no dependen de una, sino de dos, tres, y en general, de varias indeterminadas. Así, en los primeros capítulos del libro se estudiaron las formas lineales y cuadráticas, que representan ejemplos de estos polinomios. En general, se llama *polinomio* $f(x_1, x_2, \ldots, x_n)$ *en* n *indeterminadas* $x_1, x_2, \ldots, x_n$ *sobre un campo* P, a la suma de un número finito de términos de la forma $x_1^{k_1} x_2^{k_2} \ldots x_n^{k_n}$, donde $k_i \geqslant 0$, con coeficientes del campo P; se supone que el polinomio $f(x_1, x_2, \ldots, x_n)$ no contiene términos semejantes y que se consideran solamente términos con coeficientes diferentes de cero. Dos polinomios en n indeterminadas, $f(x_1, x_2, \ldots, x_n)$ y $g(x_1, x_2, \ldots, x_n)$, se consideran *iguales* (o *idénticamente iguales*), si son iguales sus coeficientes de potencias iguales.

Dado un polinomio $f(x_1, x_2, \ldots, x_n)$ sobre un campo P, se llama *grado con respecto a la indeterminada* x_i, $i = 1, 2, \ldots, n$, al exponente máximo con que figura x_i en los términos de este polinomio. Puede ocurrir eventualmente que este grado sea igual 0, lo cual significa que a pesar de que f se considere como polinomio en n indeterminadas $x_1, x_2, \ldots, x_i, \ldots, x_n$, en realidad, la indeterminada x_i no figura en su expresión.

Por otra parte, si llamamos *grado del término*

$$x_1^{k_1} x_2^{k_2} \ldots x_n^{k_n}$$

al número $k_1 + k_2 + \ldots + k_n$, o sea, a la suma de los exponentes de las indeterminadas, el *grado del polinomio* $f(x_1, x_2, \ldots, x_n)$ (o sea, el grado respecto del **conjunto de las indeterminadas**) será el grado superior de sus términos. En particular, al igual que en el caso de una indeterminada, son polinomios de grado cero solamente los elementos del campo P, diferentes de cero. Por otra parte, del mismo modo que en el caso de los polinomios en una indeterminada, el cero es el único polinomio en n indeterminadas cuyo grado está indefinido. Claro, en el caso general, un polinomio puede contener

unos cuantos términos de grado superior, por lo cual, no se puede hablar de un término superior (según el grado) del polinomio.

Para los polinomios en n indeterminadas sobre un campo P, las operaciones de sumar y multiplicar se definen del modo siguiente:

Se llama *suma* de los polinomios $f(x_1, x_2, \ldots, x_n)$ y $g(x_1, x_2, \ldots, x_n)$ al polinomio cuyos coeficientes se obtienen sumando los coeficientes correspondientes de los polinomios f y g; naturalmente, si en este caso algún término figura solamente en uno de los polinomios f, g, el coeficiente de éste en el otro polinomio se supone igual a cero. El producto de dos «monomios» se define por la igualdad:

$$ax_1^{k_1}x_2^{k_2}\ldots x_n^{k_n}\cdot bx_1^{l_1}x_2^{l_2}\ldots x_n^{l_n}=(ab)\,x_1^{k_1+l_1}x_2^{k_2+l_2}\ldots x_n^{k_n+l_n},$$

después de lo cual, el *producto* de los polinomios $f(x_1, x_2, \ldots, x_n)$ y $g(x_1, x_2, \ldots, x_n)$ se define como el resultado de la multiplicación término a término y la consiguiente reducción de términos semejantes.

Definidas las operaciones de este modo, el conjunto de los polinomios en n indeterminadas sobre el campo P se convierte en un anillo conmutativo que, además, carece de divisores de cero. En efecto, para $n = 1$ nuestras definiciones coinciden con las que se dieron en el § 20 para el caso de polinomios en una indeterminada. Supongamos que se ha demostrado que los polinomios en $n-1$ indeterminadas $x_1, x_2, \ldots, x_{n-1}$ con coeficientes del campo P forman un anillo sin divisores de cero. Todo polinomio en n indeterminadas $x_1, x_2, \ldots, x_{n-1}, x_n$ se puede representar de un modo único como un polinomio en la indeterminada x_n con coeficientes que son polinomios en $x_1, x_2, \ldots, x_{n-1}$; recíprocamente, todo polinomio en x_n con coeficientes del anillo de los polinomios en $x_1, x_2, \ldots, x_{n-1}$ sobre el campo P se puede considerar, naturalmente, como un polinomio sobre el mismo campo P en todo el conjunto de las indeterminadas $x_1, x_2, \ldots, x_{n-1}, x_n$. Se comprueba sin dificultad que la correspondencia biunívoca, obtenida entre los polinomios en n indeterminadas y los polinomios en una indeterminada sobre el anillo de los polinomios en $n-1$ indeterminadas, es un isomorfismo con respecto a las operaciones de sumar y multiplicar. La proposición que se demuestra se deduce de que los polinomios en una indeterminada sobre el anillo de los polinomios en $n-1$ indeterminadas forman ellos mismos un anillo que, además, por ser un anillo de polinomios en una indeterminada sobre un anillo sin divisores de cero, tampoco contiene divisores de cero (véase § 47).

Por consiguiente, queda demostrada la existencia del *anillo de los polinomios en n indeterminadas sobre el campo P*; este anillo se designa con la notación $P[x_1, x_2, \ldots, x_n]$.

El estudio siguiente permite examinar el anillo de los polinomios en n indeterminadas desde otro punto de vista. Supongamos que el

campo P está contenido como subanillo en un anillo conmutativo L. Tomemos en L n elementos $\alpha_1, \alpha_2, \ldots, \alpha_n$ y hallemos el subanillo mínimo L' del anillo L que contiene a estos elementos y a todo el campo P, o sea, el subanillo que se obtiene *por adjunción* de los elementos $\alpha_1, \alpha_2, \ldots, \alpha_n$ al campo P. El subanillo L' consta de todos los elementos del anillo L que se expresan mediante los elementos $\alpha_1, \alpha_2, \ldots, \alpha_n$ y los elementos del campo P aplicando la suma, resta y multiplicación. Fácilmente se observa que éstos son exactamente los elementos del anillo L que se pueden expresar (aplicando las operaciones que subsisten en el anillo L) en forma de polinomios en $\alpha_1, \alpha_2, \ldots, \alpha_n$ con coeficientes de P; además, estos elementos, como elementos del anillo L, se pueden sumar y multiplicar precisamente según las leyes de adición y multiplicación de los polinomios en n indeterminadas.

Claro, por lo general, un elemento dado β del subanillo L' puede poseer muchas expresiones distintas en forma de polinomio en $\alpha_1, \alpha_2, \ldots, \alpha_n$ con coeficientes del campo P. Si esta expresión es **única** para cualquier β de L', es decir, que diferentes polinomios en $\alpha_1, \alpha_2, \ldots, \alpha_n$ son elementos distintos del anillo L' (y, por consiguiente, del anillo L), el sistema de elementos $\alpha_1, \alpha_2, \ldots, \alpha_n$ se llama *algebraicamente independiente* sobre el campo P. En caso contrario, se llama *algebraicamente dependiente**. De aquí, se puede hacer la siguiente conclusión:

Si un campo P es un subanillo de un anillo conmutativo L y si el sistema de elementos $\alpha_1, \alpha_2, \ldots, \alpha_n$ de L es algebraicamente independiente sobre P, el subanillo L' del anillo L, engendrado por adjunción de los elementos $\alpha_1, \alpha_2, \ldots, \alpha_n$ al campo P, es isomorfo al anillo de los polinomios $P[x_1, x_2, \ldots, x_n]$.

Entre otras propiedades del anillo de los polinomios en n indeterminadas $P[x_1, x_2, \ldots, x_n]$, señalemos la siguiente: este anillo se puede incluir en el *campo de fracciones racionales $P(x_1, x_2, \ldots, x_n)$ en n indeterminadas sobre el campo P*. Todo elemento de este campo se puede expresar en la forma $\frac{f}{g}$, donde f y g son polinomios del anillo $P[x_1, x_2, \ldots x_n]$, siendo, además, $\frac{f}{g}=\frac{\varphi}{\psi}$ cuando, y sólo cuando, $f\psi = g\varphi$. La suma y el producto de estas fracciones racionales se efectúan según las leyes que, como se indicó en el § 45, se cumplen para los cocientes en cualquier campo. La demostración de la existencia del campo $P(x_1, x_2, \ldots, x_n)$ se hace igual que en el § 50 para el caso $n = 1$.

* Los conceptos correspondientes para el caso $n = 1$ fueron introducidos en el § 47: un elemento α, algebraicamente independiente sobre el campo P, en el sentido de la definición que se acaba de dar, se llamaba entonces ***trascendente*** sobre P; en el caso contrario, ***algebraico*** sobre P.

Para los polinomios en varias indeterminadas se puede construir la teoría de la divisibilidad que generaliza a la teoría de la divisibilidad de los polinomios en una indeterminada, estudiada en los cap. 5 y 10. Mas, como no entra en nuestro plan el estudio detallado del anillo de los polinomios en varias indeterminadas, nos limitaremos solamente a la cuestión de la descomposición de un polinomio en factores irreducibles.

Introduzcamos primero el siguiente concepto: si todos los términos de un polinomio $f(x_1, x_2, \ldots, x_n)$ son de un mismo grado s, éste se llama *polinomio homogéneo* o, abreviadamente, *forma* de grado s (ya conocemos las *formas lineales* y *cuadráticas*, se pueden considerar luego las formas cúbicas, todos los términos de las cuales son de grado 3 con respecto del conjunto de las indeterminadas, etc). **Todo polinomio en n indeterminadas se representa unívocamente en forma de una suma de unas cuantas formas en estas indeterminadas que son, además, de distinto grado:** para obtener la representación buscada es suficiente agrupar todos los términos de un mismo grado. Así, pues, el polinomio de cuarto grado $f(x_1, x_2, x_3) = 3x_1x_2^2 - 7x_1^2x_3^2 + x_2 - 5x_1x_2x_3 + x_1^4 - 2x_3 - 6 + x_3^3$ es la suma de la forma de cuarto grado $x_1^4 - 7x_1^2x_3^2$, de la forma cúbica $3x_1x_2^2 - 5x_1x_2x_3 + x_3^3$, de la forma lineal $x_2 - 2x_3$ y del término independiente -6 (forma de grado cero).

Demostremos ahora el siguiente **teorema:**

El grado del producto de dos polinomios en n indeterminadas, diferentes de cero, es igual a la suma de los grados de estos polinomios.

Supongamos primero que se dan dos **formas:** $\varphi(x_1, x_2, \ldots, x_n)$ de grado s y $\psi(x_1, x_2, \ldots, x_n)$ de grado t. El producto de cualquier término de la forma φ por cualquier término de la forma ψ es, evidentemente, de grado $s+t$ y, por esto, el producto $\varphi\psi$ será una forma de grado $s+t$, pues, el producto de términos semejantes no puede anular a todos los coeficientes de este producto, ya que en el anillo $P[x_1, x_2, \ldots, x_n]$ no hay divisores de cero.

Si se dan ahora unos polinomios arbitrarios $f(x_1, x_2, \ldots, x_n)$ y $g(x_1, x_2, \ldots, x_n)$ de grado s y t, respectivamente, representando cada uno de ellos en forma de una suma de formas de grados distintos, obtenemos:

$$f(x_1, x_2, \ldots, x_n) = \varphi(x_1, x_2, \ldots, x_n) + \ldots,$$
$$g(x_1, x_2, \ldots, x_n) = \psi(x_1, x_2, \ldots, x_n) + \ldots,$$

donde φ y ψ son formas de grado s y t, respectivamente, y los puntos suspensivos sustituyen a las sumas de las formas de grado menor. Entonces,

$$fg = \varphi\psi + \ldots;$$

por lo demostrado, la forma $\varphi\psi$ es de grado $s+t$, y como todos los términos sustituidos por puntos suspensivos tienen menor grado, el grado del producto fg será igual a $s+t$. El teorema queda demostrado.

El polinomio φ se llama *divisor* del polinomio f y f es *divisible* por φ, si en el anillo $P[x_1, x_2, \ldots, x_n]$ existe un polinomio ψ tal, que $f = \varphi\psi$. Fácilmente se observa que las propiedades de divisibilidad I-IX del § 21 se conservan también en el caso general que ahora estudiamos. Se dice que un polinomio f de grado k, $k \geqslant 1$, es *reducible* sobre un campo P, si se descompone en un producto de polinomios del anillo $P[x_1, x_2, \ldots, x_n]$, de grados menores que k, e *irreducible*, en caso contrario.

Todo polinomio del anillo $P[x_1, x_2, \ldots, x_n]$, de grado distinto de cero, se descompone en un producto de factores irreducibles. Esta descomposición es única, salvo factores de grado cero.

Este teorema generaliza los resultados correlativos del § 48, referentes a los polinomios en una indeterminada. Su primera tesis se demuestra repitiendo palabra por palabra los razonamientos del párrafo indicado. La demostración de la segunda tesis presenta ya dificultades considerables. Antes

de exponerla, observemos que de la segunda tesis se deduce este corolario: *si el producto de dos polinomios, f y g, del anillo $P[x_1, x_2, \ldots, x_n]$, es divisible por un polinomio irreducible p, al menos uno de estos polinomios es divisible por p.* En efecto, en caso contrario obtendríamos para el producto fg dos descomposiciones en factores irreducibles, una de las cuales no contendría a p, mientras que la otra le contendría.

Supongamos que el teorema ya está demostrado para los polinomios en n indeterminadas y queremos demostrarlo para los polinomios en $n+1$ indeterminadas, $x, x_1, x_2, \ldots, x_n$. Escribamos este polinomio en la forma $\varphi(x)$; por consiguiente, sus coeficientes serán polinomios en $x_1, x_2, \ldots, x_n$. Para estos coeficientes el teorema ya está demostrado, es decir, cada uno de ellos se descompone unívocamente en un producto de factores irreducibles. Llamemos *primitivo* (más exacto, *primitivo sobre el anillo* $P[x_1, x_2, \ldots, x_n]$) al polinomio $\varphi(x)$, si sus coeficientes no contienen ningún factor común irreducible, o sea, si éstos son primos entre sí, y demostremos el siguiente **lema de Gauss:**

El producto de dos polinomios primitivos también es un polinomio primitivo.

En efecto, sean dados los polinomios primitivos

$$f(x) = a_0x^k + a_1x^{k-1} + \ldots + a_ix^{k-i} + \ldots + a_k,$$
$$g(x) = b_0x^l + b_1x^{l-1} + \ldots + b_jx^{l-j} + \ldots + b_l$$

con coeficientes del anillo $P[x_1, x_2, \ldots, x_n]$ y sea

$$f(x)\,g(x) = c_0x^{k+l} + c_1x^{k+l-1} + \ldots + c_{i+j}x^{k+l-(i+j)} + \ldots + c_{k+l}.$$

Si este producto no es primitivo, los coeficientes $c_0, c_1, \ldots, c_{k+l}$ tienen que poseer un factor común irreducible $p = p(x_1, x_2, \ldots, x_n)$. Como no todos los coeficientes del polinomio primitivo $f(x)$ son divisibles por p, supongamos que a_i es el primer coeficiente que no es divisible por p; análogamente, designamos con b_j el primer coeficiente del polinomio $g(x)$ que no es divisible por p. Multiplicando término a término los polinomios $f(x)$ y $g(x)$ y agrupando los términos que contienen a $x^{k+l-(i+j)}$, obtenemos:

$$c_{i+j} = a_ib_j + a_{i-1}b_{j+1} + a_{i-2}b_{j+2} + \ldots + a_{i+1}b_{j-1} + a_{i+2}b_{j-2} + \ldots$$

El primer miembro de esta igualdad es divisible por el polinomio irreducible p. Sin duda, son divisibles por éste también todos los términos del segundo miembro, menos el primero; en efecto, en virtud de las condiciones impuestas a la elección de i y j, todos los coeficientes $a_{i-1}, a_{i-2}, \ldots$, y también $b_{j-1}, b_{j-2}, \ldots$, *son divisibles por p. De esto se deduce que el producto a_ib_j también* es divisible por p y, por esto, como se indicó anteriormente, tiene que ser divisible por p al menos uno de los polinomios a_i, b_j, lo cual, sin embargo, no tiene lugar. Con esto se termina la demostración del lema, suponiendo que el *teorema fundamental se verifica para los polinomios en n indeterminadas.*

Como ya sabemos, el anillo $P[x_1, x_2, \ldots, x_n]$ está contenido en el campo de fracciones racionales $P(x_1, x_2, \ldots, x_n)$ que designaremos con Q:

$$Q = P(x_1, x_2, \ldots, x_n).$$

Consideremos el anillo de los polinomios $Q[x]$. Si un polinomio $\varphi(x)$ pertenece a este anillo, cada uno de sus coeficientes se representa en forma de un *cociente de polinomios* del anillo $P[x_1, x_2, \ldots, x_n]$. *Sacando fuera de paréntesis* el común denominador de estos cocientes, y después, los factores comunes de los numeradores, se puede representar $\varphi(x)$ en la forma

$$\varphi(x) = \frac{a}{b}\,f(x).$$

Aquí, a y b son polinomios del anillo $P[x_1, x_2, \ldots, x_n]$ y $f(x)$ es un polinomio en x con coeficientes de $P[x_1, x_2, \ldots, x_n]$, que es además primitivo, pues sus coeficientes ya no tienen factores comunes.

De este modo, a cada polinomio $\varphi(x)$ del anillo $Q[x]$ se pone en correspondencia un polinomio primitivo $f(x)$. *Dado $\varphi(x)$, el polinomio $f(x)$ queda determinado unívocamente, salvo un factor de P distinto de cero.* En efecto, supongamos que

$$\varphi(x) = \frac{a}{b} f(x) = \frac{c}{d} g(x),$$

donde $g(x)$ es de nuevo un polinomio primitivo. Entonces,

$$adf(x) = bcg(x).$$

Por lo tanto, ad y bc se han obtenido sacando todos los factores comunes de los coeficientes de un mismo polinomio sobre el anillo $P[x_1, x_2, \ldots, x_n]$. Como en este anillo subsiste (por la hipótesis de inducción) el teorema de unicidad de la descomposición, de esto se deduce que ad y bc pueden diferenciarse entre sí solamente en un factor de grado cero. Por consiguiente, los polinomios primitivos $f(x)$ y $g(x)$ se diferencian entre sí en este mismo factor.

Al producto de dos polinomios del anillo $Q[x]$ le corresponde el producto de los polinomios primitivos correspondientes. En efecto, si

$$\varphi(x) = \frac{a}{b} f(x), \quad \psi(x) = \frac{c}{d} g(x)$$

donde $f(x)$ y $g(x)$ son polinomios primitivos, resulta

$$\varphi(x)\,\psi(x) = \frac{ac}{bd} f(x)\,g(x).$$

Pero, como se ha demostrado más arriba, el producto $f(x)\,g(x)$ es un polinomio primitivo.

Señalemos también que, *si el polinomio $\varphi(x)$ del anillo $Q[x]$ es irreducible sobre el campo Q, su polinomio primitivo correspondiente $f(x)$, considerado como un polinomio en $x, x_1, x_2, \ldots, x_n$, también será irreducible, y viceversa.* En efecto, si el polinomio f es reducible, $f = f_1 f_2$, ambos factores tienen que contener la indeterminada x, pues, en caso contrario, el polinomio f no sería primitivo. De aquí resulta la descomposición del polinomio $\varphi(x)$ sobre el campo Q:

$$\varphi(x) = \frac{a}{b} f(x) = \left(\frac{a}{b} f_1\right) f_2.$$

Recíprocamente, si el polinomio $\varphi(x)$ es reducible sobre Q, $\varphi(x) = \varphi_1(x)\,\varphi_2(x)$, los polinomios primitivos $f_1(x)$ y $f_2(x)$ correspondientes a $\varphi_1(x)$ y $\varphi_2(x)$ contendrán x, pero como se demostró antes, su producto es igual a $f(x)$ (salvo un factor del campo P).

Tomemos ahora un polinomio primitivo f y descompongámoslo en factores irreducibles, $f = f_1 \cdot f_2 \ldots f_k$. Todos estos factores no sólo tienen que contener a la indeterminada x, sino que tienen que ser incluso polinomios primitivos, pues, en caso contrario, el polinomio f no sería primitivo. *Esta descomposición del polinomio primitivo f es única salvo factores del campo P.* En efecto, en virtud del lema precedente, se puede considerar esta descomposición como la descomposición de $f(x)$ en factores irreducibles sobre el campo Q; mas, para los polinomios en una indeterminada sobre un campo dado, ya es conocida la unicidad de la descomposición que se verifica, salvo factores de Q; pero, en nuestro caso, como todos los factores f_i son primitivos, ésta se verifica, salvo factores de P.

Después de haber demostrado estos lemas, partiendo de la hipótesis de inducción, se hace sin dificultad alguna la demostración de nuestro teorema fundamental. En efecto, todo polinomio irreducible del anillo $P[x, x_1, x_2, \ldots, x_n]$, o es un polinomio irreducible del anillo $P[x_1, x_2, \ldots, x_n]$, o es un polinomio primitivo irreducible. De aquí se deduce que, dada una descomposición del polinomio $\varphi(x, x_1, x_2, \ldots, x_n)$ en factores irreducibles, agrupando los factores se puede representar φ en la forma

$$\varphi(x, x_1, x_2, \ldots, x_n) = a(x_1, x_2, \ldots, x_n)\, f(x, x_1, x_2, \ldots, x_n),$$

donde a no depende de x y f es un polinomio primitivo. Sin embargo, ya sabemos que esta descomposición de φ es única, salvo factores de P. Pero, por otra parte, *como, por la hipótesis de inducción, subsiste la unicidad de la descomposición en factores irreducibles para el polinomio a en n indeterminadas, y como esta unicidad está demostrada en el lema anterior para el polinomio primitivo f, queda también completamente demostrado nuestro teorema* para el caso de $n+1$ indeterminadas.

De los lemas demostrados anteriormente se deduce un corolario interesante: *si un polinomio $\varphi(x)$ con coeficientes de $P[x_1, x_2, \ldots, x_n]$ es reducible sobre el campo $Q = P(x_1, x_2, \ldots, x_n)$, entonces se puede descomponer en factores que dependen de x y cuyos coeficientes son polinomios del anillo $P[x_1, x_2, \ldots, x_n]$.* En efecto, si al polinomio $\varphi(x)$ le corresponde el polinomio primitivo $f(x)$, de modo que $\varphi(x) = af(x)$, entonces la descomposición de $\varphi(x)$ implica la descomposición de $f(x)$; pero esto último implica a su vez la descomposición de $\varphi(x)$ sobre el anillo $P[x_1, x_2, \ldots, x_n]$.

A diferencia del caso de los polinomios en una indeterminada que, como ya sabemos por el § 49, se pueden descomponer en factores lineales sobre una ampliación adecuadamente elegida del campo fundamental considerado, *existen sobre cualquier campo P polinomios de cualquier grado en varias* (dos o más) *indeterminadas que son* **absolutamente irreducibles**, o sea, polinomios que se mantienen irreducibles sobre cualquier ampliación de este campo.

De este tipo es el polinomio

$$f(x, y) = \varphi(x) + y,$$

donde $\varphi(x)$ es un polinomio arbitrario en una indeterminada sobre un campo P. En efecto, si en cierta ampliación $\overline{P}$ del campo P existiese la descomposición

$$f(x, y) = g(x, y)\, h(x, y),$$

entonces, expresando g y h según las potencias de y, obtendríamos, por ejemplo,

$$g(x, y) = a_0(x)\, y + a_1(x), \quad h(x, y) = b_0(x),$$

o sea, h no dependería de y; y como $a_0(x)b_0(x) = 1$, resultaría que $b_0(x)$ sería de grado cero y, por lo tanto, h no dependería tampoco de x.

Ordenación lexicográfica de los términos de un polinomio. Para los polinomios en una indeterminada se tienen dos métodos naturales de ordenación de los términos: según las potencias decrecientes de la indeterminada y según las potencias crecientes de la misma. En el caso de polinomios en varias indeterminadas, tales métodos no existen; por ejemplo, el polinomio de quinto grado en tres indeterminadas

$$f(x_1, x_2, x_3) = x_1x_2^2x_3^2 + x_1^4x_3 + x_2^3x_3^2 + x_1^2x_2x_3^2,$$

puede escribirse también en la forma

$$f(x_1, x_2, x_3) = x_1^4x_3 + x_1^2x_2x_3^2 + x_1x_2^2x_3^2 + x_2^3x_3^2,$$

sin que haya ningún motivo para dar preferencia a una de estas expresiones ante la otra. Pero existe, sin embargo, un método completamente determinado de ordenación de los términos de un polinomio en varias indeterminadas que depende, por cierto, de la numeración elegida de las indeterminadas; para los polinomios en una indeterminada, este método se reduce a la ordenación de los términos según las potencias decrecientes de la indeterminada. Este método, denominado *lexicográfico*, está dictado por el procedimiento común de ordenación de las palabras en los diccionarios («vocabularios»): suponiendo que las letras están ordenadas como está convenido en el alfabeto, la posición relativa en el diccionario de dos palabras dadas se determina por sus primeras letras; si éstas coinciden, por sus segundas letras, etc.

Sea dado un polinomio $f(x_1, x_2, \ldots, x_n)$ del anillo $P[x_1, x_2, \ldots, x_n]$ y dos términos distintos de él:

$$x_1^{k_1}x_2^{k_2} \ldots x_n^{k_n}, \tag{1}$$

$$x_1^{l_1}x_2^{l_2} \ldots x_n^{l_n}, \tag{2}$$

cuyos coeficientes son elementos de P, diferentes de cero. Como los términos (1) y (2) son distintos, al menos una de las diferencias de los exponentes de las indeterminadas

$$k_i - l_i, \quad i = 1, 2, \ldots, n,$$

es diferente de cero. El término (1) se considerará *superior* al término (2) (y el término (2), *inferior* al término (1)), si la primera de estas diferencias, distinta de cero, es positiva, o sea, si existe una i, $1 \leqslant i \leqslant n$, tal que

$$k_1 = l_1, \quad k_2 = l_2, \ \ldots, \ k_{i-1} = l_{i-1}, \text{ pero } k_i > l_i.$$

En otras palabras, el término (1) será superior al término (2), si el exponente de x_1 en (1) es mayor que en (2) o, siendo estos exponentes iguales, si el exponente de x_2 en (1) es mayor que en (2), etc. Por supuesto, el hecho de que el término (1) sea superior al término (2) no implica que el grado del primero con respecto al conjunto de las indeterminadas sea mayor que el del segundo. Por ejemplo, el primero de los términos

$$x_1^3x_2x_3, \quad x_1x_2^5x_3^2$$

es superior al segundo, a pesar de que es de menor grado.

Es evidente que, de dos términos distintos de un polinomio $f(x_1, x_2, \ldots, x_n)$, uno de ellos es superior al otro. Fácilmente se

comprueba también que, si el término (1) es superior al término (2), y éste, a su vez, es superior al término

$$x_1^{m_1}x_2^{m_2}\dots x_n^{m_n}, \tag{3}$$

o sea, que existe una j, $1 \leqslant j \leqslant n$, tal, que

$$l_1 = m_1,\ l_2 = m_2,\ \dots,\ l_{j-1} = m_{j-1},\ \text{pero}\ l_j > m_j,$$

el término (1) es superior al término (3), independiente de que sea i mayor, igual o menor que j. Por lo tanto, de cada dos términos, poniendo delante el que sea superior, obtenemos una ordenación determinada de los términos del polinomio $f(x_1, x_2, \dots, x_n)$, llamada lexicográfica.

Así, pues, la ordenación de los términos en el polinomio

$$f(x_1, x_2, x_3, x_4) = x_1^4 + 3x_1^2x_2^3x_3 - x_1^2x_2^3x_4^2 + 5x_1x_3x_4^2 + 2x_2 + x_3^3x_4 - 4$$

es lexicográfica.

En la expresión lexicográfica de un polinomio $f(x_1, x_2, \dots, x_n)$, uno de sus términos ocupará el primer lugar, o sea, será superior a todos los demás. Este se llama *término superior del polinomio*; en el ejemplo precedente, el término superior es x_1^4. Respecto a los términos superiores, demostraremos un **lema** que se aplicará en la demostración del teorema fundamental del siguiente párrafo:

El término superior del producto de dos polinomios en n indeterminadas es igual al producto de los términos superiores de los factores.

En efecto, supongamos que se multiplican los polinomios $f(x_1, x_2, \dots, x_n)$ y $g(x_1, x_2, \dots, x_n)$. Si

$$ax_1^{k_1}x_2^{k_2}\dots x_n^{k_n} \tag{4}$$

es el término superior del polinomio $f(x_1, x_2, \dots, x_n)$, y

$$a'x_1^{s_1}x_2^{s_2}\dots x_n^{s_n} \tag{5}$$

es otro término cualquiera del mismo, existe un valor i, $1 \leqslant i \leqslant n$, tal que

$$k_1 = s_1,\ \dots,\ k_{i-1} = s_{i-1},\quad k_i > s_i.$$

Si, por otra parte,

$$bx_1^{l_1}x_2^{l_2}\dots x_n^{l_n}, \tag{6}$$

$$b'x_2^{t_1}x_2^{t_2}\dots x_n^{t_n} \tag{7}$$

son el término superior y otro término cualquiera del polinomio $g(x_1, x_2, \dots, x_n)$, existe un valor j, $1 \leqslant j \leqslant n$, tal que

$$l_1 = t_1,\ \dots,\ l_{j-1} = t_{j-1},\ l_j > t_j.$$

Multiplicando los términos (4) y (6), y también los términos (5) y (7), obtenemos:

$$abx_1^{k_1+l_1}x_2^{k_2+l_2}\ldots x_n^{k_n+l_n}, \tag{8}$$

$$a'b'x_1^{s_1+t_1}x_2^{s_2+t_2}\ldots x_n^{s_n+t_n}. \tag{9}$$

Sin embargo, fácilmente se comprueba que el término (8) es superior al término (9); si, por ejemplo, $i \leqslant j$, resulta,

$$k_1+l_1=s_1+t_1, \ \ldots, \ k_{i-1}+l_{i-1}=s_{i-1}+t_{i-1}, \text{ pero } k_i+l_i > s_i+t_i,$$

pues $k_i > s_i$, $l_i \geqslant t_i$. Del mismo modo se comprueba que el término (8) es superior al producto de los términos (4) y (7), y superior al producto de los términos (5) y (6). Por consiguiente, el término (8), que es el producto de los términos superiores de los polinomios f y g, es superior a todos los demás términos que se obtienen multiplicando término a término los polinomios f y g, y, por lo tanto, este término no puede eliminarse al reducir los términos semejantes, o sea, se mantiene en el producto fg como término superior.

§ 52. Polinomios simétricos

Entre los polinomios en varias indeterminadas se distinguen los que no varían con cualquier permutación de las indeterminadas. Por consiguiente, en tales polinomios figuran todas las indeterminadas de un modo simétrico, por lo cual se llaman *polinomios simétricos* (o *funciones simétricas*). Los ejemplos más elementales son: la suma de todas las indeterminadas $x_1 + x_2 + \ldots + x_n$, la suma de los cuadrados de las indeterminadas $x_1^2 + x_2^2 + \ldots + x_n^2$, el producto de las indeterminadas $x_1x_2 \ldots x_n$, etc. En virtud de la posibilidad de expresar cualquier sustitución de n símbolos en forma de un producto de trasposiciones (véase el § 3), para demostrar que un polinomio es simétrico, es suficiente comprobar que éste no varía al efectuar una trasposición cualquiera de dos indeterminadas.

A continuación se estudiarán los polinomios simétricos en n indeterminadas con coeficientes de un campo P. Está claro que *la suma, diferencia y producto de dos polinomios simétricos son también simétricos*, es decir, los polinomios simétricos forman un subanillo en el anillo $P[x_1, x_2, \ldots, x_n]$ de todos los polinomios en n indeterminadas sobre el campo P, denominado *anillo de los polinomios simétricos en n indeterminadas sobre el campo P*. Todos los elementos del campo P pertenecen a este anillo (o sea, todos los polinomios de grado cero, y también el cero), ya que éstos no varían al efectuar cualquier permutación de las indeterminadas. Cualquier otro polinomio simétrico contiene indispensablemente todas las n indeterminadas, e incluso con respecto a ellas tiene un mismo grado. En efecto, si el polinomio simétrico $f(x_1, x_2, \ldots, x_n)$ tiene un término en el

que la indeterminada x_i figura con el exponente k, entonces tiene también el término que se obtiene de este último mediante la trasposición de las indeterminadas x_i y x_j, o sea, el que contiene a la indeterminada x_j con el mismo exponente k.

Los n polinomios simétricos en n indeterminadas que se exponen a continuación se llaman *polinomios simétricos elementales*:

$$\left.\begin{aligned} \sigma_1 &= x_1 + x_2 + \ldots + x_n, \\ \sigma_2 &= x_1x_2 + x_1x_3 + \ldots + x_{n-1}x_n, \\ \sigma_3 &= x_1x_2x_3 + x_1x_2x_4 + \ldots + x_{n-2}x_{n-1}x_n, \\ &\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ \sigma_{n-1} &= x_1x_2 \ldots x_{n-1} + x_1x_2 \ldots x_{n-2}x_n + \ldots + x_2x_3 \ldots x_n \\ \sigma_n &= x_1x_2 \ldots x_n. \end{aligned}\right\} \quad (1)$$

Estos polinomios que, evidentemente, son simétricos, desempeñan un papel muy importante en la teoría de los polinomios simétricos. Su origen se debe a las fórmulas de Vieta (véase el § 24). Por esto, se puede decir que *los coeficientes de un polinomio en una indeterminada, cuyo coeficiente superior es igual a la unidad, son, salvo el signo, los polinomios simétricos elementales en sus raíces.* Esta relación de los *polinomios simétricos elementales con las fórmulas* de Vieta es muy importante para las aplicaciones de los polinomios simétricos a la teoría de los polinomios en una indeterminada, y es la causa por la que ahora los estudiamos.

Como los polinomios simétricos en n indeterminadas $x_1, x_2, \ldots, x_n$ sobre el campo P forman un anillo, resultan evidentes las proposiciones siguientes: es un polinomio simétrico cualquier potencia entera y positiva de cualquiera de los polinomios elementales simétricos, y también el producto de tales potencias, tomado además con cualquier coeficiente de P y, finalmente, cualquier suma de los productos indicados. En otras palabras, *cualquier polinomio en los polinomios simétricos elementales $\sigma_1, \sigma_2, \ldots, \sigma_n$, con coeficientes de P, considerado como un polinomio en las indeterminadas $x_1, x_2, \ldots, x_n$, es simétrico.* Así, pues, pongamos $n = 3$ y tomemos el polinomio $\sigma_1\sigma_2 + 2\sigma_3$. Sustituyendo σ_1, σ_2 y σ_3 por sus expresiones, obtenemos:

$$\sigma_1\sigma_2 + 2\sigma_3 = x_1^2x_2 + x_1^2x_3 + x_1x_2^2 + x_2^2x_3 + x_1x_3^2 + x_2x_3^2 + 5x_1x_2x_3;$$

evidentemente, en el segundo miembro figura un polinomio simétrico en x_1, x_2, x_3.

Recíproco a este resultado es el siguiente **teorema fundamental de los polinomios simétricos:**

Todo polinomio simétrico en las indeterminadas $x_1, x_2, \ldots, x_n$ sobre el campo P es un polinomio en los polinomios simétricos elementales $\sigma_1, \sigma_2, \ldots, \sigma_n$ con coeficientes pertenecientes al campo P.

En efecto, sea dado un polinomio simétrico

$$f(x_1, x_2, \ldots, x_n)$$

y supongamos que en su expresión lexicográfica el término superior es

$$a_0 x_1^{k_1} x_2^{k_2} \ldots x_n^{k_n}. \tag{2}$$

Los exponentes de las indeterminadas en este término tienen que satisfacer a las desigualdades

$$k_1 \geqslant k_2 \geqslant \ldots \geqslant k_n. \tag{3}$$

En efecto, supongamos que para cierta i, $k_i < k_{i+1}$. El polinomio $f(x_1, x_2, \ldots, x_n)$, siendo simétrico, tiene que contener el término

$$a_0 x_1^{k_1} x_2^{k_2} \ldots x_i^{k_{i+1}} x_{i+1}^{k_i} \ldots x_n^{k_n}, \tag{4}$$

que se obtiene del término (2) mediante una trasposición de las indeterminadas x_i y x_{i+1}. Sin embargo, esto es absurdo, puesto que el término (4), en el sentido de la ordenación lexicográfica, es superior al término (2); en efecto, los exponentes de $x_1, x_2, \ldots, x_{i-1}$ en ambos términos coinciden, pero el exponente de x_i en el término (4) es mayor que en el término (2).

Consideremos ahora el siguiente producto de polinomios elementales simétricos (en virtud de las desigualdades (3), todos los exponentes son no negativos):

$$\varphi_1 = a_0 \sigma_1^{k_1-k_2} \sigma_2^{k_2-k_3} \ldots \sigma_{n-1}^{k_{n-1}-k_n} \sigma_n^{k_n}. \tag{5}$$

Este polinomio en las indeterminadas $x_1, x_2, \ldots, x_n$ es simétrico y su término superior es igual al término (2). En efecto, los términos superiores de los polinomios $\sigma_1, \sigma_2, \sigma_3, \ldots, \sigma_n$ son iguales a $x_1, x_1x_2, x_1x_2x_3, \ldots, x_1x_2 \ldots x_n$, respectivamente, y como al final del párrafo anterior se demostró que el término superior del producto es igual al producto de los términos superiores de los factores, el término superior del polinomio φ_1 es

$$a_0 x_1^{k_1-k_2} (x_1x_2)^{k_2-k_3} (x_1x_2x_3)^{k_3-k_4} \ldots$$
$$\ldots (x_1x_2 \ldots x_{n-1})^{k_{n-1}-k_n} (x_1x_2 \ldots x_n)^{k_n} = a_0 x_1^{k_1} x_2^{k_2} \ldots x_n^{k_n}.$$

De aquí se deduce que, al restar φ_1 de f, los términos superiores de estos polinomios se eliminan entre sí, o sea, el término superior del polinomio simétrico $f - \varphi_1 = f_1$ resulta menor que el término (2), que es el superior en el polinomio f. Repitiendo este mismo procedimiento para el polinomio f_1, cuyos coeficientes pertenecen evidentemente al campo P, llegamos a la igualdad

$$f_1 = \varphi_2 + f_2,$$

donde φ_2 es un producto de potencias de polinomios simétricos elementales con cierto coeficiente del campo P, y f_2, un polinomio *simétrico cuyo término superior es inferior al término superior de* f_1. De aquí, resulta la igualdad

$$f = \varphi_1 + \varphi_2 + f_2.$$

Continuando este proceso, para cierto s obtenemos $f_s = 0$. De este modo, llegaremos a obtener para f una expresión en forma de un polinomio en $\sigma_1, \sigma_2, \ldots, \sigma_n$ con coeficientes de P:

$$f(x_1, x_2, \ldots, x_n) = \sum_{i=1}^{s} \varphi_i = \varphi(\sigma_1, \sigma_2, \ldots, \sigma_n).$$

En efecto, si este proceso fuese indefinido *, obtendríamos una *sucesión indefinida* de polinomios simétricos

$$f_1, f_2, \ldots, f_s, \ldots, \tag{6}$$

donde el término superior de cada uno de ellos sería inferior a los términos superiores de los precedentes polinomios y, por lo tanto, inferior a (2). Pero, si

$$bx_1^{l_1}x_2^{l_2} \ldots x_n^{l_n} \tag{7}$$

es el término superior del polinomio f_s, como este último es simétrico, resultan las desigualdades

$$l_1 \geqslant l_2 \geqslant \ldots \geqslant l_n, \tag{8}$$

semejantes a las desigualdades (3). Por otra parte, como el término (2) es superior al término (7), se tiene

$$k_1 \geqslant l_1. \tag{9}$$

Además, se observa fácilmente que los sistemas de números enteros *no negativos* $l_1, l_2, \ldots, l_n$, *que satisfacen a las desigualdades* (8) y (9), se pueden elegir solamente de un número finito de modos. En efecto, incluso cuando no se insiste en el cumplimiento de la condición (8), si se supone solamente que todas las l_i, $i = 1, 2, \ldots, n$, no son mayores que k_1, resulta ya que los números l_i se pueden elegir solamente de $(k_1 + 1)^n$ modos. De aquí se deduce que la sucesión de polinomios (6) con los términos superiores estrictamente decrecientes, no puede ser indefinida.

El teorema queda demostrado.

De la relación entre los polinomios elementales simétricos y las fórmulas de Vieta, indicadas anteriormente, se desprende el siguiente

* Hay que tener presente que, por lo general, el polinomio φ_s contiene también términos que no existen en el polinomio f_{s-1} y, por esto, el paso de f_{s-1} a $f_s = f_{s-1} - \varphi_s$ no sólo está ligado con la eliminación de ciertos términos de f_{s-1}, sino también con la aparición de nuevos términos. Aquí $s = 1, 2, \ldots$

corolario importante del teorema fundamental de los polinomios simétricos:

Sea $f(x)$ un polinomio en una indeterminada sobre el campo P, con el coeficiente superior igual a la unidad. Entonces, cualquier polinomio simétrico (con coeficientes de P) en las **raíces** *del polinomio $f(x)$, pertenecientes a un campo de descomposición del polinomio $f(x)$ sobre P, es un polinomio (con coeficientes de P) en los* **coeficientes** *del polinomio $f(x)$ y, por lo tanto, es un elemento del campo P.*

La demostración expuesta del teorema fundamental proporciona a la vez un método para la averiguación práctica de las expresiones de los polinomios simétricos mediante los polinomios elementales. Hagamos primero la siguiente notación: siendo

$$ax_1^{k_1}x_2^{k_2}\ldots x_n^{k_n} \tag{10}$$

un producto de potencias de las indeterminadas $x_1, x_2, \ldots, x_n$ (algunos de los exponentes pueden ser iguales a cero), mediante

$$S\,(ax_1^{k_1}x_2^{k_2}\ldots x_n^{k_n}) \tag{11}$$

designaremos la suma de todos los términos que se obtienen de (10) al permutar las indeterminadas de todos los modos posibles. Evidentemente, éste es un polinomio simétrico y homogéneo. También es evidente que cualquier polinomio simétrico en n indeterminadas que contenga al término (10), contiene también todos los demás términos del polinomio (11). Por ejemplo, $S(x_1)=\sigma_1$, $S(x_1x_2)=\sigma_2$, $S(x_1^2)$ es la suma de los cuadrados de todas las indeterminadas, etc.

Ejemplo. Expresar el polinomio simétrico $f=S(x_1^2x_2)$ en n indeterminadas mediante los polinomios simétricos elementales.

Aquí, el término superior es $x_1^2x_2$, y por esto, $\varphi_1=\sigma_1^{2-1}\sigma_2=\sigma_1\sigma_2$, o sea,

$$\varphi_1=(x_1+x_2+\ldots+x_n)(x_1x_2+x_1x_3+\ldots+x_{n-1}x_n)=$$
$$=S(x_1^2x_2)+3S(x_1x_2x_3),$$

de donde

$$f_1=f-\varphi_1=-3S(x_1x_2x_3)=-3\sigma_3.$$

Por esto $f=\varphi_1+f_1=\sigma_1\sigma_2-3\sigma_3$.

En ejemplos más complicados es conveniente determinar primero qué términos pueden figurar en la expresión del polinomio dado mediante los polinomios elementales y hallar después los coeficientes de estos términos por el método de los coeficientes indeterminados.

Ejemplos. 1. Hallar la expresión para el polinomio simétrico $f=S(x_1^2x_2^2)$.

Ya sabemos (véase la demostración del teorema fundamental) que los términos del polinomio buscado $\varphi(\sigma_1, \sigma_2, \ldots, \sigma_n)$ se determinan mediante los términos superiores de los polinomios simétricos $f_1, f_2, \ldots$, siendo inferiores estos términos al término superior del polinomio dado f, o sea, inferiores a $x_1^2x_2^2$. Hallemos todos los productos $x_1^{l_1}x_2^{l_2}\ldots x_n^{l_n}$ que satisfacen a las condiciones siguientes: 1) son inferiores al término $x_1^2x_2^2$; 2) pueden servir de términos superiores para los polinomios simétricos, o sea, satisfacen a las desigualdades $l_1\geqslant l_2\geqslant \ldots \geqslant l_n$; 3) son de cuarto grado con respecto al conjunto de las indeterminadas (pues, como ya sabemos, todos los polinomios $f_1, f_2, \ldots$ tienen el mismo grado que el polinomio homogéneo f). Escribiendo solamente las combinaciones correspondientes de los exponentes e indicando al lado los productos de las

potencias de σ determinados por ellos, obtenemos la tabla siguiente:

$$22\,000 \ldots \sigma_1^{2-2}\sigma_2^{2-0} = \sigma_2^2,$$
$$21\,100 \ldots \sigma_1^{2-1}\sigma_2^{1-1}\sigma_3^{1-0} = \sigma_1\sigma_3,$$
$$11\,110 \ldots \sigma_1^{1-1}\sigma_2^{1-1}\sigma_3^{1-1}\sigma_4^{1-0} = \sigma_4.$$

Por lo tanto, el polinomio f tiene la forma

$$f = \sigma_2^2 + A\sigma_1\sigma_3 + B\sigma_4.$$

Hemos hecho el coeficiente de σ_2 igual a la unidad, pues, este término se determina por el término superior del polinomio f que, como ya sabemos por la demostración del teorema fundamental, tiene este mismo coeficiente. Los coeficientes A y B los hallaremos del modo siguiente.

Pongamos $x_1 = x_2 = x_3 = 1$, $x_4 = \ldots = x_n = 0$. Fácilmente se observa que, para estos valores de las indeterminadas, el polinomio f toma el valor 3 y los polinomios σ_1, σ_2, σ_3 y σ_4, los valores 3, 3, 1 y 0, respectivamente. Por esto

$$3 = 9 + A\cdot 3\cdot 1 + B\cdot 0,$$

de donde $A = -2$. Pongamos ahora $x_1 = x_2 = x_3 = x_4 = 1$, $x_5 = \ldots = x_n = 0$. Los valores de los polinomios f, σ_1, σ_2, σ_3 y σ_4 son 6, 4, 6, 4, 1, respectivamente. Por esto,

$$6 = 36 - 2\cdot 4\cdot 4 + B\cdot 1,$$

de donde $B = 2$. Por lo tanto, la expresión buscada para f es

$$f = \sigma_2^2 - 2\sigma_1\sigma_3 + 2\sigma_4.$$

2. Hallar la suma de los cubos de las raíces del polinomio

$$f(x) = x^4 + x^3 + 2x^2 + x + 1.$$

Para la resolución de este problema, hallemos la expresión mediante los polinomios simétricos elementales para el polinomio simétrico $S(x_1^3)$. Aplicando el mismo método que en el ejemplo anterior, obtenemos la tabla

$$3\,000 \ldots \sigma_1^3,$$
$$2\,100 \ldots \sigma_1\sigma_2,$$
$$1\,110 \ldots \sigma_3,$$

y, por esto,

$$S(x_1^3) = \sigma_1^3 + A\sigma_1\sigma_2 + B\sigma_3.$$

Poniendo primero $x_1 = x_2 = 1$, $x_3 = \ldots = x_n = 0$, y después, $x_1 = x_2 = x_3 = 1$, $x_4 = \ldots = x_n = 0$, obtenemos, $A = -3$, $B = 3$, o sea,

$$S(x_1^3) = \sigma_1^3 - 3\sigma_1\sigma_2 + 3\sigma_3. \qquad (12)$$

Para hallar la suma de los cubos de las raíces del polinomio $f(x)$ dado, en virtud de las fórmulas de Víeta, en la expresión que hemos hallado hay que sustituir σ_1 por el coeficiente de x^3 con signo contrario, o sea, por -1; σ_2, por el coeficiente de x^2, o sea, por 2; y, por fin, σ_3, por el coeficiente de x con signo contrario, o sea, por -1. Por consiguiente, la suma de los cubos de las raíces que nos interesa es igual a

$$(-1)^3 - 3\cdot(-1)\cdot 2 + 3\cdot(-1) = 2.$$

El lector puede comprobar este resultado teniendo en cuenta que las raíces de $f(x)$ son: $i, -i, -\frac{1}{2} + i\frac{\sqrt{3}}{2}$ y $-\frac{1}{2} - i\frac{\sqrt{3}}{2}$. Está claro también que la fórmula (12) no depende del polinomio $f(x)$ dado y permite hallar la suma de los cubos de las raíces de cualquier polinomio.

El método obtenido en la demostración del teorema fundamental para expresar un polinomio simétrico f mediante los polinomios elementales, conduce a un polinomio en $\sigma_1, \sigma_2, \ldots, \sigma_n$ completamente determinado. Resulta que de ningún modo se puede obtener para f otra expresión distinta mediante $\sigma_1, \sigma_2, \ldots, \sigma_n$. Esto muestra el siguiente **teorema de unicidad**:

Todo polinomio simétrico posee una expresión única en forma de un polinomio en los polinomios simétricos elementales.

Demostremos este teorema. Si un polinomio simétrico $f(x_1, x_2, \ldots, x_n)$ sobre el campo P poseyese dos expresiones distintas mediante $\sigma_1, \sigma_2, \ldots, \sigma_n$:

$$f(x_1, x_2, \ldots, x_n) = \varphi(\sigma_1, \sigma_2, \ldots, \sigma_n) = \psi(\sigma_1, \sigma_2, \ldots, \sigma_n),$$

la diferencia

$$\chi(\sigma_1, \sigma_2, \ldots, \sigma_n) = \varphi(\sigma_1, \sigma_2, \ldots, \sigma_n) - \psi(\sigma_1, \sigma_2, \ldots, \sigma_n)$$

sería un polinomio en $\sigma_1, \sigma_2, \ldots, \sigma_n$, diferente de cero, es decir, que no todos sus coeficientes serían iguales a cero, mientras que la sustitución en este polinomio de $\sigma_1, \sigma_2, \ldots, \sigma_n$ por sus expresiones mediante $x_1, x_2, \ldots, x_n$ nos daría el cero del anillo $P[x_1, x_2, \ldots, x_n]$. Por esto, no queda más que demostrar que, si un polinomio $\chi(\sigma_1, \sigma_2, \ldots, \sigma_n)$ es diferente de cero, o sea, que tiene por lo menos un coeficiente diferente de cero, el polinomio $g(x_1, x_2, \ldots, x_n)$, obtenido de χ sustituyendo $\sigma_1, \sigma_2, \ldots, \sigma_n$ por sus expresiones mediante $x_1, x_2, \ldots, x_n$:

$$\chi(\sigma_1, \sigma_2, \ldots, \sigma_n) = g(x_1, x_2, \ldots, x_n), \tag{13}$$

es también diferente de cero.

Si $a\sigma_1^{k_1}\sigma_2^{k_2} \ldots \sigma_n^{k_n}$ es uno de los términos del polinomio χ, siendo $a \neq 0$, entonces, como ya sabemos por la demostración del teorema fundamental, después de sustituir todas las σ por sus expresiones (1), obtenemos un polinomio en $x_1, x_2, \ldots, x_n$, cuyo término superior (en el sentido de la ordenación lexicográfica) es

$$ax_1^{k_1}(x_1x_2)^{k_2} \ldots (x_1x_2 \ldots x_n)^{k_n} = ax_1^{l_1}x_2^{l_2} \ldots x_n^{l_n},$$

donde

$$\begin{aligned} l_1 &= k_1 + k_2 + \ldots + k_n, \\ l_2 &= \quad\;\; k_2 + \ldots + k_n, \\ &\ldots\ldots\ldots\ldots\ldots \\ l_n &= \qquad\qquad\quad k_n. \end{aligned}$$

De aquí resulta,

$$k_i = l_i - l_{i+1}, \;\; k_n = l_n, \quad i = 1, 2, \ldots, n-1,$$

o sea, conociendo los exponentes $l_1, l_2, \ldots, l_n$ se pueden restituir los exponentes $k_1, k_2, \ldots, k_n$ del término inicial del polinomio χ. Por lo tanto, distintos términos del polinomio χ, considerados como polinomios en $x_1, x_2, \ldots, x_n$, tienen términos superiores **diferentes**.

Consideremos ahora todos los términos del polinomio χ; para cada uno de ellos, hallemos el término superior de su expresión en forma de un polinomio en $x_1, x_2, \ldots, x_n$ y entre estos términos superiores elijamos el que sea **superior** en el sentido de la ordenación lexicográfica. Como ya se advirtió antes, este término no tiene semejantes entre los términos superiores que se obtienen de los demás términos del polinomio χ y, como por la condición, este término es superior a cada uno de estos términos **superiores**, es superior, por consiguiente, a todos los demás términos que se obtienen sustituyendo los elementos $\sigma_1, \sigma_2, \ldots, \sigma_n$ en los términos del polinomio χ por sus expresiones (1). Por lo tanto, hemos hallado un término que, al pasar de $\chi(\sigma_1, \sigma_2, \ldots, \sigma_n)$ a $g(x_1, x_2, \ldots, x_n)$, aparece (con un coeficiente diferente de cero) una sola vez, por lo cual, no puede simplificarse con ninguno. De aquí se deduce que no todos los coeficientes del polinomio $g(x_1, x_2, \ldots, x_n)$ son iguales a cero, o sea, que este polinomio no es el cero del anillo $P[x_1, x_2, \ldots, x_n]$, como se quería demostrar.

Es evidente que el teorema demostrado se puede enunciar también del modo siguiente:

El sistema de los polinomios simétricos elementales $\sigma_1, \sigma_2, \ldots, \sigma_n$, considerados como elementos del anillo de los polinomios $P[x_1, x_2, \ldots, x_n]$, es algebraicamente independiente sobre el campo P.

§ 53. Observaciones complementarias sobre los polinomios simétricos

Observaciones sobre el teorema fundamental. La demostración del teorema fundamental de los polinomios simétricos expuesta en el párrafo anterior, permite hacer algunos complementos esenciales al enunciado del teorema, los cuales se aplicarán más adelante. Ante todo, los coeficientes del polinomio $\varphi(\sigma_1, \sigma_2, \ldots, \sigma_n)$, hallado como expresión del polinomio simétrico $f(x_1, x_2, \ldots, x_n)$ mediante los polinomios simétricos elementales, no sólo pertenecen al campo P, sino que *se obtienen incluso de los coeficientes del polinomio f aplicando las operaciones de adición y sustracción, o sea, pertenecen al anillo L engendrado por los coeficientes del polinomio f dentro del campo P.*

En efecto, como fácilmente se observa, todos los coeficientes del polinomio φ_1 (véase la fórmula (5) del párrafo precedente) son, con respecto a las indeterminadas $x_1, x_2, \ldots, x_n$, múltiplos enteros

del coeficiente a_0 del término superior del polinomio f y, por lo tanto, pertenecen al anillo L. Supongamos que ya está demostrado que pertenecen a L todos los coeficientes (con respecto a $x_1, x_2, \ldots, x_n$) de los polinomios $\varphi_1, \varphi_2, \ldots, \varphi_l$, entonces, los coeficientes del polinomio $f_l = f - \varphi_1 - \varphi_2 - \ldots - \varphi_l$ también pertenecen a L y, por ende, pertenecen también a L todos los coeficientes del polinomio φ_{l+1} con respecto a $x_1, x_2, \ldots, x_n$.

Por otra parte, *el grado del polinomio* $\varphi(\sigma_1, \sigma_2, \ldots, \sigma_n)$ *con respecto al conjunto* $\sigma_1, \sigma_2, \ldots, \sigma_n$ *es igual al grado que tiene el polinomio* $f(x_1, x_2, \ldots, x_n)$ *con respecto a cada una de las indeterminadas* x_i. En efecto, como, por el párrafo anterior, (2) es el término superior del polinomio f, k_1 es el grado de f con respecto a x_1 y por esto, en virtud de la simetría, es también el grado de f con respecto a cualquiera otra de las indeterminadas x_i. Mas, por la igualdad (5) del párrafo anterior, el grado de φ_1 con respecto al conjunto σ es igual al número

$$(k_1 - k_2) + (k_2 - k_3) + \ldots + (k_{n-1} - k_n) + k_n = k_1.$$

Por otra parte, como el término superior del polinomio f_1 es inferior al término superior del polinomio f, el grado de f_1 con respecto a cada x_i no será superior al grado de f con respecto a cada una de estas indeterminadas. Pero el polinomio φ_2 desempeña para f_1 el mismo papel que φ_1 para f, por consiguiente, el grado de φ_2 con respecto al conjunto σ es igual al grado de f_1 con respecto a cada x_i, o sea, no es mayor que k_1, etc. Por lo tanto, el grado de $\varphi(\sigma_1, \sigma_2, \ldots, \sigma_n)$ tampoco es mayor que k_1. Pero, como ninguna φ_i con $i > 1$, puede contener todas las $\sigma_1, \sigma_2, \ldots, \sigma_n$ elevadas a las mismas potencias que φ_1, el grado de $\varphi(\sigma_1, \sigma_2, \ldots, \sigma_n)$ es exactamente igual a k_1. Con esto, nuestra proposición queda demostrada.

Sea, finalmente, $a\sigma_1^{l_1}\sigma_2^{l_2} \ldots \sigma_n^{l_n}$ uno de los términos del polinomio $\varphi(\sigma_1, \sigma_2, \ldots, \sigma_n)$. Llamemos *peso* de este término al número

$$l_1 + 2l_2 + \ldots + nl_n,$$

o sea, a la suma de los exponentes multiplicados por los índices que corresponden a σ_i. En otras palabras, como se deduce del teorema del grado de un producto de polinomios, demostrado en el § 51, el peso es el grado del término que consideramos con respecto al conjunto de las indeterminadas $x_1, x_2, \ldots, x_n$. Entonces, se verifica la siguiente proposición:

Si un polinomio simétrico homogéneo $f(x_1, x_2, \ldots, x_n)$ *es de grado s con respecto al conjunto de las indeterminadas, todos los términos de su expresión* $\varphi(\sigma_1, \sigma_2, \ldots, \sigma_n)$ *mediante* σ *tienen un mismo peso, igual a s.*

En efecto, si (2) del párrafo anterior es el término superior del polinomio homogéneo f, se tiene

$$s = k_1 + k_2 + \ldots + k_n.$$

Mas, el peso del término φ_1, según (5) del párrafo precedente, es igual a

$$(k_1 - k_2) + 2(k_2 - k_3) + \ldots + (n-1)(k_{n-1} - k_n) + nk_n =$$
$$= k_1 + k_2 + k_3 + \ldots + k_n,$$

o sea, también es igual a s. Luego, el polinomio $f_1 = f - \varphi_1$, como diferencia de dos polinomios homogéneos de grado s, también es un polinomio homogéneo de grado s y, por esto, el peso del término φ_2 del polinomio φ también es igual a s, etc.

Fracciones racionales simétricas. El *teorema fundamental* de los polinomios simétricos se puede generalizar para el caso de fracciones racionales. Llamemos *simétrica* a la fracción racional $\frac{f}{g}$ en n indeterminadas $x_1, x_2, \ldots, x_n$, si se mantiene igual a sí misma al hacer cualquier permutación de las indeterminadas. Fácilmente se demuestra que *esta definición no depende de que se tome la fracción* $\frac{f}{g}$ *o una fracción* $\frac{f_0}{g_0}$ *equivalente a ella*. En efecto, si ω es una permutación de las indeterminadas y φ es un polinomio arbitrario en estas indeterminadas, convengamos en designar con φ^ω el polinomio en que se transforma φ al efectuar la permutación ω. Según la hipótesis, para cualquier ω, se tiene,

$$\frac{f}{g} = \frac{f^\omega}{g^\omega},$$

o sea, $fg^\omega = gf^\omega$. Por otra parte, de la igualdad

$$\frac{f}{g} = \frac{f_0}{g_0}$$

resulta, $fg_0 = gf_0$, de donde $f^\omega g_0^\omega = g^\omega f_0^\omega$. Multiplicando por f ambos miembros de la última igualdad, obtenemos:

$$ff^\omega g_0^\omega = fg^\omega f_0^\omega = gf^\omega f_0^\omega,$$

de donde, después de simplificar por f^ω, resulta: $fg_0^\omega = gf_0^\omega$, o sea,

$$\frac{f_0^\omega}{g_0^\omega} = \frac{f}{g} = \frac{f_0}{g_0}.$$

Se verifica el siguiente **teorema**:

Toda fracción racional simétrica en las indeterminadas $x_1, x_2, \ldots$ $\ldots, x_n$ con coeficientes del campo P, se expresa en forma de una

fracción racional en los polinomios simétricos elementales $\sigma_1, \sigma_2, \ldots, \sigma_n$, con coeficientes pertenecientes de nuevo a P.

En efecto, sea dada una fracción racional simétrica

$$\frac{f(x_1, x_2, \ldots, x_n)}{g(x_1, x_2, \ldots, x_n)}.$$

Suponiendo que ésta es irreducible, se podría demostrar que, tanto f como g, son polinomios simétricos. Sin embargo, el camino que se sigue a continuación es el más sencillo. Si el polinomio g no es simétrico, multiplicamos el numerador y el denominador por el producto de todos los $n! - 1$ polinomios que se obtienen de g efectuando todas las sustituciones posibles no idénticas de las indeterminadas. Fácilmente se comprueba que ahora el denominador es un polinomio simétrico. En virtud de la simetría de toda la fracción, de aquí se deduce ahora que el numerador es también simétrico y, por esto, para la demostración del teorema no queda más que expresar el numerador y denominador mediante los polinomios simétricos elementales.

Sumas de potencias. En las aplicaciones aparecen frecuentemente los polinomios simétricos $s_k = x_1^k + x_2^k + \ldots + x_n^k$, $k = 1, 2, \ldots$, o sea, las sumas de las potencias k-ésimas de las indeterminadas $x_1, x_2, \ldots, x_n$. Estos polinomios, llamados *sumas de potencias*, tienen que expresarse mediante los polinomios simétricos elementales, según el teorema fundamental. Sin embargo, es bastante difícil encontrar estas expresiones para valores grandes de k y, por lo tanto, ofrece interés la relación existente entre los polinomios $s_1, s_2, \ldots$ y $\sigma_1, \sigma_2, \ldots, \sigma_n$, que se va a establecer ahora.

En primer lugar, $s_1 = \sigma_1$. Por otra parte, siendo $k \leqslant n$, fácilmente se comprueba que se verifican las igualdades:

$$\left.\begin{array}{l} s_{k-1}\sigma_1 = s_k + S(x_1^{k-1}x_2)^{*}, \\ s_{k-2}\sigma_2 = S(x_1^{k-1}x_2) + S(x_1^{k-2}x_2x_3), \\ \ldots\ldots\ldots\ldots\ldots\ldots \\ s_{k-i}\sigma_i = S(x_1^{k-i+1}x_2\ldots x_i) + S(x_1^{k-i}x_2\ldots x_i x_{i+1}),\ 2 \leqslant i \leqslant k-2, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ s_1\sigma_{k-1} = S(x_1^2 x_2 \ldots x_{k-1}) = k\sigma_k. \end{array}\right\} \quad (1)$$

Tomando la suma alternada de estas igualdades (o sea, la suma con los signos alternados), y pasando después todos los términos a una parte de la igualdad, obtenemos la fórmula siguiente:

$$s_k - s_{k-1}\sigma_1 + s_{k-2}\sigma_2 - \ldots + (-1)^{k-1} s_1\sigma_{k-1} + (-1)^k k\sigma_k = 0 \ (k \leqslant n). \quad (2)$$

* Véase (11) del párrafo precedente.

Si $k > n$, el sistema de igualdades (1) toma la forma:

$$s_{k-1}\sigma_1 = s_k + S(x_1^{k-1}x_2),$$
$$s_{k-2}\sigma_2 = S(x_1^{k-1}x_3) + S(x_1^{k-2}x_2x_3),$$
$$\dots\dots\dots\dots\dots$$
$$s_{k-i}\sigma_i = S(x_1^{k-i+1}x_2\dots x_i) + S(x_1^{k-i}x_2\dots x_ix_{i+1}), \qquad 2 \leqslant i \leqslant n-1,$$
$$\dots\dots\dots\dots\dots$$
$$s_{k-n}\sigma_n = S(x_1^{k-n+1}x_2\dots x_n),$$

de donde se deduce la fórmula

$$s_k - s_{k-1}\sigma_1 + s_{k-2}\sigma_2 - \dots + (-1)^n s_{k-n}\sigma_n = 0 \quad (k > n). \qquad (3)$$

Las fórmulas (2) y (3) se llaman *fórmulas de Newton*. Estas ligan a las sumas de potencias con los polinomios simétricos elementales y, por consiguiente, permiten hallar sucesivamente las expresiones de $s_1, s_2, s_3, \dots$ mediante $\sigma_1, \sigma_2, \dots, \sigma_n$. Así, pues, ya sabemos que $s_1 = \sigma_1$, lo cual se deduce también de la fórmula (2). Si $k = 2 \leqslant n$ entonces, en virtud de (2), se tiene $s_2 - s_1\sigma_1 + 2\sigma_2 = 0$, de donde

$$s_2 = \sigma_1^2 - 2\sigma_2.$$

Si $k = 3 \leqslant n$, se tiene $s_3 - s_2\sigma_1 + s_1\sigma_2 - 3\sigma_3 = 0$, de donde, aplicando las expresiones ya obtenidas para s_1 y s_2, obtenemos:

$$s_3 = \sigma_1^3 - 3\sigma_1\sigma_2 + 3\sigma_3,$$

lo cual ya conocemos (véase (12) del párrafo precedente). Si $k = 3$, pero $n = 2$, por (3) se tiene $s_3 - s_2\sigma_1 + s_1\sigma_2 = 0$, de donde $s_3 = \sigma_1^3 - 3\sigma_1\sigma_2$. Aplicando las fórmulas de Newton, se puede obtener una fórmula general que exprese s_k mediante $\sigma_1, \sigma_2, \dots, \sigma_n$. Pero, debido a la complejidad de esta fórmula, omitimos su exposición.

Si el campo fundamental P es de característica 0 y, por lo tanto, tiene sentido la división por cualquier número natural n*, la fórmula (2) permite expresar sucesivamente los polinomios simétricos elementales $\sigma_1, \sigma_2, \dots, \sigma_n$, mediante las primeras n sumas de potencias $s_1, s_2, \dots, s_n$. Así, pues, $\sigma_1 = s_1$, y, por esto,

$$\sigma_2 = \frac{1}{2}(s_1\sigma_1 - s_2) = \frac{1}{2}(s_1^2 - s_2),$$

$$\sigma_3 = \frac{1}{3}(s_3 - s_2\sigma_1 + s_1\sigma_2) = \frac{1}{6}(s_1^3 - 3s_1s_2 + 2s_3)$$

etc. De aquí, y del teorema fundamental, se desprende el siguiente resultado:

* En un campo de característica p, la expresión $\frac{a}{p}$ carece de sentido si $a \neq 0$, pues, en este campo, para cualquier x, se tiene $px = 0$.

Todo polinomio simétrico en n indeterminadas $x_1, x_2, \ldots, x_n$, *sobre un campo P de característica cero, se puede expresar en forma de un polinomio en las sumas de potencias* $s_1, s_2, \ldots, s_n$ *con coeficientes pertenecientes al campo P.*

Polinomios simétricos con respecto a dos sistemas de indeterminadas. En el siguiente párrafo, así como en el § 58, se va a utilizar una generalización del concepto de polinomio simétrico. Sean dados dos sistemas de indeterminadas, $x_1, x_2, \ldots, x_n$ e $y_1, y_2, \ldots, y_r$, donde su unión

$$x_1, x_2, \ldots, x_n, y_1, y_2, \ldots, y_r \tag{4}$$

es algebraicamente independiente sobre el campo P. Un polinomio $f(x_1, x_2, \ldots, x_n, y_1, y_2, \ldots, y_r)$ sobre el campo P se llama *simétrico con respecto a los dos sistemas de indeterminadas*, si no varía al hacer cualesquiera permutaciones de las indeterminadas $x_1, x_2, \ldots, x_n$ entre sí y de las indeterminadas $y_1, y_2, \ldots, y_r$ entre sí. Si para los polinomios simétricos elementales en $x_1, x_2, \ldots, x_n$ conservamos las notaciones $\sigma_1, \sigma_2, \ldots, \sigma_n$, y designamos con $\tau_1, \tau_2, \ldots, \tau_r$, los polinomios simétricos elementales en $y_1, y_2, \ldots, y_r$, el teorema fundamental se generaliza del modo siguiente:

Todo polinomio $f(x_1, x_2, \ldots, x_n, y_1, y_2, \ldots, y_r)$ *sobre el campo P, que es simétrico con respecto a los sistemas de indeterminadas* $x_1, x_2, \ldots, x_n$ *e* $y_1, y_2, \ldots, y_r$, *se expresa en forma de un polinomio (con coeficientes de P) en los polinomios simétricos elementales respecto de estos dos sistemas de indeterminadas:*

$$f(x_1, x_2, \ldots, x_n, y_1, y_2, \ldots, y_r) = \varphi(\sigma_1, \sigma_2, \ldots, \sigma_n, \tau_1, \tau_2, \ldots, \tau_r).$$

En efecto, el polinomio f se puede considerar como un polinomio $\bar{f}(y_1, y_2, \ldots, y_r)$ de coeficientes que son polinomios en $x_1, x_2, \ldots, x_n$. Como f no varía al permutar las indeterminadas $x_1, x_2, \ldots, x_n$, los coeficientes del polinomio $\bar{f}$ serán polinomios simétricos en $x_1, x_2, \ldots, x_n$, por lo cual, en virtud del teorema fundamental, se expresan en forma de polinomios (con coeficientes de P) en $\sigma_1, \sigma_2, \ldots, \sigma_n$. Por otra parte, el polinomio $\bar{f}(y_1, y_2, \ldots, y_r)$, considerado sobre el campo $P(x_1, x_2, \ldots, x_n)$, es simétrico con respecto a $y_1, y_2, \ldots, y_r$, por lo cual, se expresa en forma de un polinomio $\bar{\varphi}(\tau_1, \tau_2, \ldots, \tau_r)$. Como se ha mostrado al principio del presente párrafo, los coeficientes del polinomio $\bar{\varphi}$ se expresan mediante los coeficientes del polinomio $\bar{f}$ mediante la suma y la resta y, por consiguiente, también son polinomios en $\sigma_1, \sigma_2, \ldots, \sigma_n$. Evidentemente, esto nos conduce a la expresión buscada de f mediante $\sigma_1, \sigma_2, \ldots, \sigma_n, \tau_1, \tau_2, \ldots, \tau_r$.

Ejemplo. El polinomio

$$f(x_1, x_2, x_3, y_1, y_2) = x_1x_2x_3 - x_1x_2y_1 - x_1x_2y_2 - x_1x_3y_1 -$$
$$- x_1x_3y_2 - x_2x_3y_1 - x_2x_3y_2 + x_1y_1y_2 + x_2y_1y_2 + x_3y_1y_2$$

es simétrico con respecto a las indeterminadas x_1, x_2, x_3, así como con respecto a las indeterminadas y_1, y_2, pero no es simétrico con respecto al conjunto de todas las cinco indeterminadas, lo cual se observa trasponiendo las indeterminadas x_1 e y_1. Hallemos la expresión de f mediante $\sigma_1, \sigma_2, \sigma_3, \tau_1, \tau_2$:

$$f = x_1x_2x_3 - (x_1x_2 + x_1x_3 + x_2x_3)\, y_1 - (x_1x_2 + x_1x_3 + x_2x_3)\, y_2 +$$
$$+ (x_1 + x_2 + x_3)\, y_1y_2 = \sigma_3 - \sigma_2y_1 - \sigma_2y_2 + \sigma_1y_1y_2 = \sigma_3 - \sigma_2\tau_1 + \sigma_1\tau_2.$$

Naturalmente, el teorema que se acaba de demostrar se generaliza también al caso de tres y de un número mayor de sistemas de indeterminadas.

Para los polinomios que son simétricos con respecto a dos sistemas de indeterminadas se verifica también el **teorema de unicidad** de la representación mediante los polinomios simétricos elementales. En otras palabras, se verifica el siguiente **teorema:**

El sistema unido

$$\sigma_1,\ \sigma_2,\ \ldots,\ \sigma_n,\ \tau_1,\ \tau_2,\ \ldots,\ \tau_r$$

de polinomios simétricos elementales en los sistemas dados de indeterminadas $x_1, x_2, \ldots, x_n$ *e* $y_1, y_2, \ldots, y_r$, *es algebraicamente independiente sobre el campo* P.

En efecto, supongamos que existe un polinomio

$$\varphi(\sigma_1, \sigma_2, \ldots, \sigma_n, \tau_1, \tau_2, \ldots, \tau_r)$$

sobre el campo P, que es igual a cero, a pesar de que no todos sus coeficientes son iguales a cero. Este polinomio se puede considerar como un polinomio $\psi(\tau_1, \tau_2, \ldots, \tau_r)$ de coeficientes que son polinomios en $\sigma_1, \sigma_2, \ldots, \sigma_n$. Por consiguiente, se puede suponer que ψ es un polinomio en $\tau_1, \tau_2, \ldots, \tau_r$ sobre el campo de fracciones racionales:

$$Q = P(x_1, x_2, \ldots, x_n).$$

El sistema $y_1, y_2, \ldots, y_r$ se mantiene algebraicamente independiente sobre el campo Q, pues, si para este sistema existiese una dependencia algebraica con coeficientes de Q, eliminando los denominadores obtendríamos una dependencia algebraica en el sistema (4), en contra de la hipótesis. Basándose en el teorema de unicidad del párrafo anterior, resulta ahora que el sistema $\tau_1, \tau_2, \ldots, \tau_r$ también tiene que ser algebraicamente independiente sobre el campo Q y, por esto, todos los coeficientes del polinomio ψ son iguales a cero. Pero, estos coeficientes son polinomios en $\sigma_1, \sigma_2, \ldots, \sigma_n$, por lo cual, de nuevo, en virtud del teorema de unicidad para el caso de un sistema de indeterminadas (esta vez, para el sistema

$x_1, x_2, \ldots, x_n$), los mismos coeficientes de estos últimos polinomios son iguales a cero. Con esto queda demostrado que, en contra de la hipótesis, todos los coeficientes del polinomio φ tienen que ser iguales a cero.

§ 54. Resultante. Eliminación de una indeterminada. Discriminante

Dado un polinomio $f(x_1, x_2, \ldots, x_n)$ del anillo $P[x_1, x_2, \ldots, x_n]$, se llama *solución* del mismo a un sistema de valores de las indeterminadas

$$x_1 = \alpha_1, \; x_2 = \alpha_2, \; \ldots, \; x_n = \alpha_n,$$

tomados en el campo P o en alguna ampliación $\overline{P}$ de este campo, que convierte en cero al polinomio f:

$$f(\alpha_1, \alpha_2, \ldots, \alpha_n) = 0.$$

Todo polinomio f, cuyo grado sea mayor que cero, posee soluciones. En efecto, si la indeterminada x_1 figura en la expresión de este polinomio, entonces, se pueden tomar por $\alpha_2, \ldots, \alpha_n$ los elementos arbitrarios del campo P, con la condición solamente de que el grado del polinomio $f(x_1, \alpha_2, \ldots, \alpha_n)$ se mantenga estrictamente positivo, y después, aplicando el teorema de existencia de la raíz (§ 49), se puede tomar una ampliación $\overline{P}$ del campo P, en la que el polinomio $f(x_1, \alpha_2, \ldots, \alpha_n)$ en una indeterminada x_1 tenga una raíz α_1. A la vez, observamos que la propiedad (de los polinomios de grado n en una indeterminada) de poseer en cualquier campo no más de n raíces, no se cumple para los polinomios en varias indeterminadas.

Dados unos cuantos polinomios en n indeterminadas, se puede plantear el problema del cálculo de las soluciones que son comunes a todos ellos, o sea, de las soluciones del sistema de ecuaciones que resulta al igualar a cero los polinomios dados. En el segundo capítulo se estudió detalladamente un caso particular de este problema, precisamente, el caso de sistemas de ecuaciones lineales. Sin embargo, en el caso particular inverso de una ecuación en una indeterminada, pero de grado arbitrario, no sabemos nada sobre las raíces, a excepción de que éstas existen en cierta ampliación del campo fundamental. Naturalmente, la búsqueda y el estudio de las soluciones de un sistema no lineal de ecuaciones en varias indeterminadas es un problema más complicado que, por cierto, está fuera de los márgenes de nuestro curso y es el objeto de una rama de las matemáticas, denominada geometría algebraica. Aquí nos limitaremos solamente al caso de un sistema de **dos** ecuaciones de grado arbitrario en **dos** indeterminadas y demostraremos que éste se puede reducir al caso de **una** ecuación en **una** indeterminada.

Ocupémonos primero del problema de la existencia de raíces comunes de dos polinomios en una indeterminada. Sean dados los polinomios

$$\left.\begin{array}{l} f(x) = a_0x^n + a_1x^{n-1} + \ldots + a_{n-1}x + a_n, \\ g(x) = b_0x^s + b_1x^{s-1} + \ldots + b_{s-1}x + b_s \end{array}\right\} \quad (1)$$

sobre el campo P, siendo $a_0 \neq 0$, $b_0 \neq 0$.

De los resultados del párrafo precedente, sin dificultad alguna se deduce que *los polinomios $f(x)$ y $g(x)$ poseen raíz común en cierta ampliación del campo P cuando, y sólo cuando, éstos no son primos entre sí*. Por lo tanto, el problema de la existencia de raíces comunes para los polinomios dados se puede resolver aplicándoles el algoritmo de Euclides.

Ahora señalaremos otro método para dar una respuesta a este problema. Sea $\overline{P}$ una ampliación tal del campo P, en la que $f(x)$ tenga n raíces $\alpha_1, \alpha_2, \ldots, \alpha_n$, y $g(x)$ tenga s raíces, $\beta_1, \beta_2, \ldots, \beta_s$; por $\overline{P}$ se puede tomar el campo de descomposición del producto $f(x)\, g(x)$. El elemento

$$R(f, g) = a_0^s b_0^n \prod_{i=1}^{n} \prod_{j=1}^{s} (\alpha_i - \beta_j) \quad (2)$$

del campo $\overline{P}$ se llama *resultante* de los polinomios $f(x)$ y $g(x)$. Es evidente que *$f(x)$ y $g(x)$ poseen en $\overline{P}$ raíz común cuando, y sólo cuando*, $R(f, g) = 0$. Como

$$g(x) = b_0 \prod_{j=1}^{s} (x - \beta_j)$$

se tiene,

$$g(\alpha_i) = b_0 \prod_{j=1}^{s} (\alpha_i - \beta_j);$$

la resultante $R(f, g)$ se puede expresar también en la forma

$$R(f, g) = a_0^s \prod_{i=1}^{n} g(\alpha_i). \quad (3)$$

En la definición de la resultante, los polinomios $f(x)$ y $g(x)$ no se emplean de un modo simétrico. En efecto,

$$R(g, f) = b_0^n a_0^s \prod_{j=1}^{s} \prod_{i=1}^{n} (\beta_j - \alpha_i) = (-1)^{ns} R(f, g). \quad (4)$$

En correspondencia con (3), $R(g, f)$ se puede expresar en la forma

$$R(g, f) = b_0^n \prod_{j=1}^{s} f(\beta_j). \quad (5)$$

La expresión (2) para la resultante exige conocer las raíces de los polinomios $f(x)$ y $g(x)$ y, por esto, prácticamente es inútil para la resolución del problema de la existencia de una raíz común de estos polinomios. Sin embargo, resulta que *la resultante* $R(f, g)$ *se puede expresar en forma de un polinomio en los coeficientes* $a_0, a_1, \ldots, a_n, b_0, b_1, \ldots, b_s$ *de los polinomios* $f(x)$ *y* $g(x)$.

La posibilidad de tal representación se deduce fácilmente de los resultados del párrafo anterior. En efecto, la fórmula (2) muestra que la resultante $R(f, g)$ es un polinomio simétrico en dos sistemas de indeterminadas: en el sistema $\alpha_1, \alpha_2, \ldots, \alpha_n$ y en el sistema $\beta_1, \beta_2, \ldots, \beta_s$. Por esto, como se demostró al fin del párrafo anterior, ésta se representa en forma de un polinomio en los polinomios simétricos elementales en estos dos sistemas de indeterminadas, o sea, en virtud de las fórmulas de Vieta, en forma de un polinomio en los cocientes $\frac{a_i}{a_0}$, $i = 1, 2, \ldots, n$, y $\frac{b_j}{b_0}$, $j = 1, 2, \ldots, s$; el factor $a_0^s b_0^n$, incluido en (2), libra de a_0 y b_0 al denominador de la expresión obtenida. Por cierto, sería muy difícil hallar la expresión de la resultante mediante los coeficientes con los métodos expuestos en los párrafos anteriores, por lo que emplearemos otro método.

La expresión que hallaremos para la resultante de los polinomios (1) será válida para cualquier par de estos polinomios. Precisando, se supondrá que el **sistema de raíces**

$$\alpha_1, \alpha_2, \ldots, \alpha_n, \beta_1, \beta_2, \ldots, \beta_s \tag{6}$$

de los polinomios (1) **es un sistema de** $n + s$ **indeterminadas independientes, o sea, es un sistema de** $n + s$ **elementos, algebraicamente independiente sobre el campo** P en el sentido del § 51.

Obtendremos una expresión para la resultante que, considerada como un polinomio en las indeterminadas (6) (después de sustituir los coeficientes mediante las raíces por las fórmulas de Vieta), será también igual al segundo miembro de la igualdad (2), considerado también como un polinomio en las indeterminadas (6).

Entendiendo la igualdad precisamente en el sentido de identidad con respecto al sistema de las indeterminadas (6), demostraremos que *la resultante* $R(f, g)$ *de los polinomios* (1) *es igual al siguiente determinante de orden* $n + s$:

$$D = \begin{vmatrix} a_0 & a_1 & \ldots & a_n & & & \\ & a_0 & a_1 & \ldots & a_n & & \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ & & & a_0 & a_1 & \ldots & a_n \\ b_0 & b_1 & \ldots & b_s & & & \\ & b_0 & b_1 & \ldots & b_s & & \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ & & & b_0 & b_1 & \ldots & b_s \end{vmatrix} \begin{matrix} \left.\vphantom{\begin{matrix}a\\a\\a\\a\end{matrix}}\right\} s \text{ filas} \\ \left.\vphantom{\begin{matrix}a\\a\\a\\a\end{matrix}}\right\} n \text{ filas} \end{matrix} \tag{7}$$

(en los lugares libres figuran ceros). La estructura de este determinante está suficientemente clara; señalemos solamente que en su diagonal principal figura s veces el coeficiente a_0 y, después, n veces el coeficiente b_s.

Para la demostración de nuestra afirmación, calcularemos de dos modos el producto $a_0^s b_0^n DM$, donde M es el siguiente determinante auxiliar de orden $n+s$:

$$M = \begin{vmatrix} \beta_1^{n+s-1} & \beta_2^{n+s-1} & \dots & \beta_s^{n+s-1} & \alpha_1^{n+s-1} & \alpha_2^{n+s-1} & \dots & \alpha_n^{n+s-1} \\ \beta_1^{n+s-2} & \beta_2^{n+s-2} & \dots & \beta_s^{n+s-2} & \alpha_1^{n+s-2} & \alpha_2^{n+s-2} & \dots & \alpha_n^{n+s-2} \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \beta_1^2 & \beta_2^2 & \dots & \beta_s^2 & \alpha_1^2 & \alpha_2^2 & \dots & \alpha_n^2 \\ \beta_1 & \beta_2 & \dots & \beta_s & \alpha_1 & \alpha_2 & \dots & \alpha_n \\ 1 & 1 & \dots & 1 & 1 & 1 & \dots & 1 \end{vmatrix}.$$

M es el determinante de Vandermonde y, por esto, como se indicó en el § 6, es igual al producto de las diferencias de los elementos de su penúltima fila, donde, de cada elemento precedente se resta cualquier elemento posterior. Por lo tanto,

$$M = \prod_{1\leqslant i<j\leqslant s} (\beta_i-\beta_j)\cdot \prod_{j=1}^{s}\prod_{i=1}^{n} (\beta_j-\alpha_i)\cdot \prod_{1\leqslant i<j\leqslant n} (\alpha_i-\alpha_j).$$

de donde, en virtud de (4),

$$a_0^s b_0^n DM = D\cdot R(g, f)\cdot \prod_{1\leqslant i<j\leqslant s} (\beta_i-\beta_j)\cdot \prod_{1\leqslant i<j\leqslant n} (\alpha_i-\alpha_j). \qquad (8)$$

Por otra parte, calculemos el producto DM basándonos en el teorema del determinante del producto de las matrices. Multiplicando las matrices correspondientes y teniendo en cuenta que todas las α son raíces de $f(x)$ y todas las β son raíces de $g(x)$, obtenemos:

$$a_0^s b_0^n DM =$$

$$= \begin{vmatrix} \beta_1^{s-1}f(\beta_1) & \beta_2^{s-1}f(\beta_2) & \dots & \beta_s^{s-1}f(\beta_s) & 0 & 0 & \dots & 0 \\ \beta_1^{s-2}f(\beta_1) & \beta_2^{s-2}f(\beta_2) & \dots & \beta_s^{s-2}f(\beta_s) & 0 & 0 & \dots & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \beta_1 f(\beta_1) & \beta_2 f(\beta_2) & \dots & \beta_s f(\beta_s) & 0 & 0 & \dots & 0 \\ f(\beta_1) & f(\beta_2) & \dots & f(\beta_s) & 0 & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 & \alpha_1^{n-1}g(\alpha_1) & \alpha_2^{n-1}g(\alpha_2) & \dots & \alpha_n^{n-1}g(\alpha_n) \\ 0 & 0 & \dots & 0 & \alpha_1^{n-2}g(\alpha_1) & \alpha_2^{n-2}g(\alpha_2) & \dots & \alpha_n^{n-2}g(\alpha_n) \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & \dots & 0 & \alpha_1 g(\alpha_1) & \alpha_2 g(\alpha_2) & \dots & \alpha_n g(\alpha_n) \\ 0 & 0 & \dots & 0 & g(\alpha_1) & g(\alpha_2) & \dots & g(\alpha_n) \end{vmatrix}$$

Aplicando el teorema de Laplace, sacando después los factores comunes de las columnas de los determinantes y calculando los determinantes que quedan como determinantes de Vandermonde, resulta:

$$a_0^s b_0^n DM = a_0^s b_0^n \prod_{j=1}^{s} f(\beta_s) \cdot \prod_{1 \leq i < j \leq s} (\beta_i - \beta_j) \cdot \prod_{i=1}^{n} g(\alpha_i) \cdot \prod_{1 \leq i < j \leq n} (\alpha_i - \alpha_j),$$

o bien, aplicando (3) y (5),

$$a_0^s b_0^n DM = R(f, g)\, R(g, f) \cdot \prod_{1 \leq i < j \leq s} (\beta_i - \beta_j) \cdot \prod_{1 \leq i < j \leq s} (\alpha_i - \alpha_j). \qquad (9)$$

Ha resultado que los segundos miembros de las igualdades (8) y (9), considerados como polinomios en las indeterminadas (6), son iguales entre sí. Ambos miembros de la igualdad obtenida se pueden simplificar por sus factores comunes, que no son idénticamente iguales a cero. El factor común $R(g, f)$ no es igual a cero. En efecto, como por la hipótesis, $a_0 \neq 0$ y $b_0 \neq 0$, es suficiente elegir para las indeterminadas (6) valores que no sean iguales entre sí (en el campo fundamental o en alguna ampliación del mismo), para obtener en (4) un valor diferente de cero del polinomio $R(g, f)$. Del mismo modo se demuestra que los otros dos factores comunes son diferentes de cero. Simplificando por todos estos factores comunes, llegamos a la igualdad:

$$R(f, g) = D \qquad (10)$$

como se quería demostrar.

Desistamos ahora de la condición de que los coeficientes superiores de los polinomios (1) **sean diferentes de cero *.** Por consiguiente, acerca de los grados verdaderos de estos polinomios solamente se puede afirmar que éstos no son superiores a sus grados «formales» n y s, respectivamente. Ahora, la expresión (2) para la resultante carece de sentido, pues, posiblemente, los polinomios considerados tienen una cantidad de raíces menor que n o s. Por otra parte, ahora también se puede escribir el determinante (7) y como ya está demostrado que, siendo $a_0 \neq 0$, $b_0 \neq 0$, este determinante es igual a la resultante, le llamaremos también, en el caso general, *resultante* de los polinomios $f(x)$ y $g(x)$, designándole con la notación $R(f, g)$.

Pero ya no se puede asegurar que la igualdad a cero de la resultante es equivalente a la existencia de una raíz común de nuestros

---

* El hecho de que por ahora nos neguemos de la condición que habíamos impuesto al coeficiente superior del polinomio, se debe a las aplicaciones ulteriores, puesto que queremos estudiar los sistemas de polinomios en dos indeterminadas, refiriendo una de éstas a los coeficientes. Por consiguiente, el coeficiente superior puede anularse para valores particulares de esta indeterminada.

polinomios. En efecto, si $a_0 = 0$ y $b_0 = 0$, resulta que $R(f, g) = 0$, independientemente de que tengan los polinomios f y g raíces comunes o no. Sin embargo, éste es el único caso en que de la igualdad a cero de la resultante no se puede hacer la conclusión de que existen *raíces comunes de estos polinomios**. *Precisando, se verifica el* siguiente teorema:

Dados los polinomios (1) *con cualesquiera coeficientes superiores, su resultante es igual a cero cuando, y sólo cuando, estos polinomios tienen una raíz común, o bien, cuando ambos coeficientes superiores son iguales a cero.*

Demostración. El caso en que $a_0 \neq 0$, $b_0 \neq 0$, ya se estudió anteriormente y el caso en que $a_0 = b_0 = 0$ se tiene en cuenta en el enunciado del teorema. No queda más que considerar el caso en que uno de los coeficientes superiores de los polinomios (1), por ejemplo a_0, *es diferente de cero, mientras que* b_0 *es igual a cero.*

Si $b_i = 0$ para todos los i, $i = 0, 1, \ldots, s$, entonces $R(f, g) = 0$, pues el determinante (7) contiene filas que constan de ceros. Pero, entonces el polinomio $g(x)$ será idénticamente igual a cero, por lo cual, tendrá raíces comunes con $f(x)$. Si

$$b_0 = b_1 = \ldots = b_{k-1} = 0, \text{ pero } b_k \neq 0, \ k \leqslant s,$$

y

$$\overline{g}(x) = b_k x^{s-k} + b_{k+1} x^{s-k-1} + \ldots + b_{s-1} x + b_s,$$

entonces, sustituyendo por ceros los elementos $b_0, b_1, \ldots, b_{k-1}$ en el determinante (7), y aplicando el teorema de Laplace, obtenemos, evidentemente, la igualdad:

$$R(f, g) = a_0^k R(f, \overline{g}). \tag{11}$$

Sin embargo, como los coeficientes superiores de ambos polinomios f y $\overline{g}$ son diferentes de cero, por lo demostrado anteriormente, la igualdad $R(f, \overline{g}) = 0$ es condición necesaria y suficiente para la existencia de una raíz común de los polinomios f y $\overline{g}$. Por otra parte, en virtud de (11), las igualdades $R(f, g) = 0$ y $R(f, \overline{g}) = 0$ son equivalentes, y como los polinomios g y $\overline{g}$ tienen raíces iguales, obtenemos que, en el caso considerado, la igualdad a cero de la resultante $R(f, g)$ es equivalente a la existencia de una raíz común de los polinomios $f(x)$ y $g(x)$. Con esto, el teorema queda demostrado.

---

* Naturalmente, el determinante (7) también es igual a cero cuando $a_n = b_s = 0$. Mas, en este caso, los polinomios (1) tienen la raíz común 0.

Hallemos la resultante de los dos polinomios cuadrados

$$f(x) = a_0x^2 + a_1x + a_2, \quad g(x) = b_0x^2 + b_1x + b_2.$$

Según (7)

$$R(f, g) = \begin{vmatrix} a_0 & a_1 & a_2 & 0 \\ 0 & a_0 & a_1 & a_2 \\ b_0 & b_1 & b_2 & 0 \\ 0 & b_0 & b_1 & b_2 \end{vmatrix},$$

o, calculando el determinante, desarrollándolo para esto por la primera y tercera filas,

$$R(f, g) = (a_0b_2 - a_2b_0)^2 - (a_0b_1 - a_1b_0)(a_1b_2 - a_2b_1). \tag{12}$$

Así, pues, dados los polinomios

$$f(x) = x^2 - 6x + 2, \quad g(x) = x^2 + x + 5,$$

en virtud de (12), se tiene, $R(f, g) = 233$, y por esto, estos polinomios no tienen raíces comunes. Dados los polinomios

$$f(x) = x^2 - 4x - 5, \quad g(x) = x^2 - 7x + 10,$$

se tiene, $R(f, g) = 0$, o sea, estos polinomios tienen una raíz común e igual a 5.

Eliminación de una indeterminada en un sistema de dos ecuaciones con dos indeterminadas. Sean dados dos polinomios f y g en dos indeterminadas x e y, con coeficientes pertenecientes a un campo P. Escribiremos estos polinomios según las potencias decrecientes de la indeterminada x:

$$\left.\begin{aligned} f(x, y) &= a_0(y)x^k + a_1(y)x^{k-1} + \ldots + a_{k-1}(y)x + a_k(y), \\ g(x, y) &= b_0(y)x^l + b_1(y)x^{l-1} + \ldots + b_{l-1}(y)x + b_l(y); \end{aligned}\right\} \tag{13}$$

los coeficientes son polinomios del anillo $P[y]$. Hallemos la resultante de los polinomios f y g, considerados como polinomios en x, y designémosla mediante $R_x(f, g)$; en virtud de (7), ésta es un polinomio en una indeterminada y, con coeficientes del campo P:

$$R_x(f, g) = F(y). \tag{14}$$

Supongamos que el sistema de polinomios (13) posee una solución común $x = \alpha$, $y = \beta$ en una ampliación del campo P. Poniendo en (13), en lugar de y el valor β, obtenemos dos polinomios, $f(x, \beta)$ y $g(x, \beta)$, en una indeterminada x. Estos polinomios tienen una raíz común α y, por consiguiente, su resultante, que en virtud de (14), es igual a $F(\beta)$, tiene que ser igual a cero, o sea, β *tiene que ser raíz de la resultante* $R_x(f, g)$. Recíprocamente, si la resultante $R_x(f, g)$ de los polinomios (13) tiene una raíz β, la resultante de los polinomios $f(x, \beta)$ y $g(x, \beta)$ es igual a cero, o sea, *o bien estos polinomios tienen una raíz común, o bien sus coeficientes superiores son iguales a cero,*

$$a_0(\beta) = b_0(\beta) = 0.$$

De este modo, el cálculo de las soluciones comunes del sistema de polinomios (13) se ha reducido al cálculo de las raíces de un polinomio (14) en una indeterminada y, o sea, como está convenido decir, *se ha eliminado la indeterminada x en el sistema de polinomios* (13).

El teorema que sigue responde a la pregunta sobre el grado del polinomio que se obtiene al eliminar una indeterminada en un sistema de dos polinomios en dos indeterminadas:

Si los polinomios $f(x, y)$ y $g(x, y)$ tienen con respecto al conjunto de las indeterminadas los grados n y s, respectivamente, el grado del polinomio $R_x(f, g)$ con respecto a la indeterminada y no es mayor que el producto ns, naturalmente, si este polinomio no es igual a cero idénticamente.

Ante todo, si se consideran dos polinomios en una indeterminada con los coeficientes superiores iguales a la unidad, según (2) su resultante $R(f, g)$ es un polinomio homogéneo en $\alpha_1, \alpha_2, \ldots, \alpha_n, \beta_1, \beta_2, \ldots, \beta_s$, de grado ns. De aquí se deduce que, si en la expresión de la resultante mediante los coeficientes $a_1, a_2, \ldots, a_n, b_1, b_2, \ldots, b_s$ figura el término

$$a_1^{k_1}a_2^{k_2}\ldots a_n^{k_n}b_1^{l_1}b_2^{l_2}\ldots b_s^{l_s}$$

y si el número

$$k_1+2k_2+\ldots+nk_n+l_1+2l_2+\ldots+sl_s$$

lo denominamos *peso* de este término, *todos los términos de la expresión de $R(f, g)$ mediante los coeficientes tienen un mismo peso, igual a ns.* Esta proposición es verídica también en el caso general para los términos de la resultante (7), si se llama *peso* del término $a_0^{k_0}a_1^{k_1}\ldots a_n^{k_n}b_0^{l_0}b_1^{l_1}\ldots b_s^{l_s}$ al número

$$0\cdot k_0+1\cdot k_1+\ldots+nk_n+0\cdot l_0+1\cdot l_1+\ldots+sl_s. \tag{15}$$

En efecto, sustituyendo en los términos del determinante (7) los factores a_0 y b_0 por la unidad, llegamos al caso ya considerado, pero los exponentes de estos factores figuran en (15) con los coeficientes 0.

Escribamos ahora los polinomios f y g en la forma siguiente:

$$f(x, y)=a_0(y)\,x^n+a_1(y)\,x^{n-1}+\ldots+a_n(y),$$
$$g(x, y)=b_0(y)\,x^s+b_1(y)\,x^{s-1}+\ldots+b_s(y).$$

Como n es el grado de $f(x, y)$ con respecto al conjunto de las indeterminadas, el grado del coeficiente $a_r(y)$, $r=0, 1, 2, \ldots, n$, no puede ser mayor que su índice r; esto mismo es cierto también para $b_r(y)$. De aquí se deduce, que el grado de cada término de la resultante $R_x(f, g)$ no es mayor que el peso de este término, o sea, no es mayor que el número ns, como se quería demostrar.

Ejemplos.

1. Hallar las soluciones del sistema de polinomios

$$f(x, y)=x^2y+3xy+2y+3,$$
$$g(x, y)=2xy-2x+2y+3.$$

Eliminemos la indeterminada x en este sistema, para lo cual, lo escribimos en la forma:

$$\left.\begin{aligned} f(x, y)&=y\cdot x^2+(3y)\cdot x+(2y+3),\\ g(x, y)&=(2y-2)\,x+(2y+3);\end{aligned}\right\} \tag{16}$$

entonces,

$$R_x(f, g) = \begin{vmatrix} y & 3y & 2y+3 \\ 2y-2 & 2y+3 & 0 \\ 0 & 2y-2 & 2y+3 \end{vmatrix} = 2y^2 + 11y + 12.$$

Las raíces de la resultante son: $\beta_1 = -4$, $\beta_2 = -\frac{3}{2}$. Para estos valores de la indeterminada y, los coeficientes superiores de los polinomios (16) no se anulan y, por esto, cada uno de ellos, junto con cierto valor de x, forma una solución del sistema dado de polinomios. Los polinomios

$$f(x, -4) = -4x^2 - 12x - 5,$$
$$g(x, -4) = -10x - 5$$

tienen una raíz común, $\alpha_1 = -\frac{1}{2}$. Los polinomios

$$f\left(x, -\frac{3}{2}\right) = -\frac{3}{2}x^2 - \frac{9}{2}x,$$
$$g\left(x, -\frac{3}{2}\right) = -5x$$

tienen una raíz común $\alpha_2 = 0$. Por lo tanto, el sistema dado de polinomios tiene dos soluciones:

$$\alpha_1 = -\frac{1}{2}, \quad \beta_1 = -4 \quad \text{y} \quad \alpha_2 = 0, \quad \beta_2 = -\frac{3}{2}.$$

2. Eliminar una indeterminada en el sistema de polinomios:

$$f(x, y) = 2x^3y - xy^2 + x + 5,$$
$$g(x, y) = x^2y^2 + 2xy^2 - 5y + 1.$$

Como estos dos polinomios son de grado 2 con respecto a la indeterminada y, mientras que uno de ellos es de grado 3 con respecto a la indeterminada x, conviene eliminar la y. Escribamos el sistema en la forma

$$\left.\begin{aligned} f(x, y) &= (-x)\cdot y^2 + (2x^3)\cdot y + (x+5), \\ g(x, y) &= (x^2+2x)\, y^2 - 5y + 1 \end{aligned}\right\} \tag{17}$$

y hallemos su resultante, aplicando la fórmula (12):

$$R_y(f, g) = [(-x)\cdot 1 - (x+5)(x^2+2x)]^2 -$$
$$- [(-x)(-5) - 2x^3(x^2+2x)]\,[2x^3\cdot 1 - (x+5)(-5)] =$$
$$= 4x^8 + 8x^7 + 11x^6 + 84x^5 + 161x^4 + 154x^3 + 96x^2 - 125x.$$

Una de las raíces de la resultante es igual a 0. Sin embargo, para este valor de la indeterminada x, ambos coeficientes superiores de los polinomios (17) se convierten en cero, y, además, como fácilmente se observa, los polinomios $f(0, y)$ y $g(0, y)$ no tienen raíces comunes. No conocemos un método para hallar las otras raíces de la resultante. Solamente se puede afirmar que si las hallásemos (por ejemplo, en el campo de descomposición de $R_y(f, g)$), ninguna de ellas anularía a ambos coeficientes superiores de los polinomios (17) y, por esto, cada una de estas raíces, junto con cierto valor de y (con uno, e incluso con varios) formaría una solución del sistema dado de polinomios.

Existen métodos que permiten eliminar sucesivamente las indeterminadas en un sistema con un número arbitrario de polinomios e indeterminadas. Pero estos métodos son demasiado complicados, por lo cual, no pueden ser incluidos en nuestro curso.

Discriminante. *Por analogía con el problema que nos ha llevado* al concepto de resultante, se puede plantear la cuestión sobre las condiciones según las cuales un polinomio $f(x)$ de grado n del anillo $P[x]$ posee raíces múltiples. Sea

$$f(x) = a_0x^n + a_1x^{n-1} + \ldots + a_{n-1}x + a_n, \quad a_0 \neq 0,$$

y supongamos que en cierta ampliación del campo P este polinomio tiene las raíces $\alpha_1, \alpha_2, \ldots, \alpha_n$. Evidentemente, *entre estas raíces hay iguales cuando, y sólo cuando, es igual a cero el producto*

$$\begin{aligned}\Delta = (\alpha_2-\alpha_1)(\alpha_3-\alpha_1)\ldots(\alpha_n-\alpha_1)\times\\ \times(\alpha_3-\alpha_2)(\alpha_4-\alpha_2)\ldots(\alpha_n-\alpha_2)\times\\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\\ \times(\alpha_n-\alpha_{n-1}) = \prod_{n\geq i>j\geq 1}(\alpha_i-\alpha_j)\end{aligned}$$

o, lo que es lo mismo, si es igual a cero el producto

$$D = a_0^{2n-2}\prod_{n\geq i>j\geq 1}(\alpha_i-\alpha_j)^2,$$

denominado *discriminante* del polinomio $f(x)$.

A diferencia del producto Δ, que puede cambiar de signo al permutar las raíces, *el discriminante D es simétrico con respecto a* $\alpha_1, \alpha_2, \ldots, \alpha_n$ y, por esto, se puede expresar mediante los coeficientes del polinomio $f(x)$. Para hallar esta expresión, **suponiendo que la característica del campo P es igual a cero**, se puede utilizar la relación existente entre el discriminante del polinomio $f(x)$ y la resultante de este polinomio y su derivada. Es natural esperar la existencia de tal relación, pues, como ya sabemos por el § 49, un polinomio tiene raíces múltiples cuando, y sólo cuando, tiene raíces comunes con su derivada $f'(x)$, por lo cual, $D = 0$ cuando, y sólo cuando, $R(f, f') = 0$.

Por la fórmula (3) del presente párrafo, se tiene:

$$R(f, f') = a_0^{n-1}\prod_{i=1}^{n} f'(\alpha_i).$$

Derivando la igualdad

$$f(x) = a_0\prod_{k=1}^{n}(x-\alpha_k),$$

resulta:

$$f'(x) = a_0 \sum_{k=1}^{n} \prod_{j \neq k} (x - \alpha_j).$$

Después de poner aquí α_i en lugar de x, todos los sumandos, a excepción del i-ésimo, se anulan, por lo cual,

$$f'(\alpha_i) = a_0 \prod_{j \neq i} (\alpha_i - \alpha_j),$$

de donde

$$R(f, f') = a_0^{n-1} \cdot a_0^n \prod_{i=1}^{n} \prod_{j \neq i} (\alpha_i - \alpha_j).$$

En este producto, para cualesquiera i y j, $i > j$, figuran dos factores: $\alpha_i - \alpha_j$ y $\alpha_j - \alpha_i$. El producto de éstos es igual a $(-1) \cdot (\alpha_i - \alpha_j)^2$, y como existen $\frac{n(n-1)}{2}$ pares de índices i, j, que satisfacen a las desigualdades $n \geqslant i > j \geqslant 1$, resulta:

$$R(f, f') = (-1)^{\frac{n(n-1)}{2}} a_0^{2n-1} \prod_{n \geqslant i > j \geqslant 1} (\alpha_i - \alpha_j)^2 = (-1)^{\frac{n(n-1)}{2}} a_0 D.$$

Ejemplo. Hallemos el discriminante del trinomio cuadrático

$$f(x) = ax^2 + bx + c.$$

Como $f'(x) = 2ax + b$, se tiene,

$$R(f, f') = \begin{vmatrix} a & b & c \\ 2a & b & 0 \\ 0 & 2a & b \end{vmatrix} = a(-b^2 + 4ac).$$

En el caso considerado, $\frac{n(n-1)}{2} = 1$, por lo cual,

$$D = -a^{-1} R(f, f') = b^2 - 4ac.$$

Esto coincide con lo que en el álgebra escolar llaman ordinariamente discriminante de la ecuación cuadrática.

Otro método para hallar el discriminante consiste en lo siguiente. Formemos el determinante de Vandermonde de las potencias de las raíces $\alpha_1, \alpha_2, \ldots, \alpha_n$. Como se demostró en el § 6,

$$\begin{vmatrix} 1 & 1 & \ldots & 1 \\ \alpha_1 & \alpha_2 & \ldots & \alpha_n \\ \alpha_1^2 & \alpha_2^2 & \ldots & \alpha_n^2 \\ \ldots & \ldots & \ldots & \ldots \\ \alpha_1^{n-1} & \alpha_2^{n-1} & \ldots & \alpha_n^{n-1} \end{vmatrix} = \prod_{n \geqslant i > j \geqslant 1} (\alpha_i - \alpha_j) = \Delta,$$

y por esto, el discriminante es igual al cuadrado de este determinante multiplicado por a_0^{2n-2}. Multiplicando este determinante por su traspuesto según la regla de multiplicación de las matrices y recordando las sumas de potencias, definidas en el párrafo precedente, resulta:

$$D = a_0^{2n-2} \begin{vmatrix} n & s_1 & s_2 & \dots & s_{n-1} \\ s_1 & s_2 & s_3 & \dots & s_n \\ s_2 & s_3 & s_4 & \dots & s_{n+1} \\ \dots & \dots & \dots & \dots & \dots \\ s_{n-1} & s_n & s_{n+1} & \dots & s_{2n-2} \end{vmatrix}, \tag{18}$$

donde s_k es la suma de las k-ésimas potencias de las raíces $\alpha_1, \alpha_2, \dots, \alpha_n$.

Ejemplo. Hallemos el discriminante del polinomio cúbico $f(x) = x^3 + ax^2 + bx + c$. Por (18), se tiene,

$$D = \begin{vmatrix} 3 & s_1 & s_2 \\ s_1 & s_2 & s_3 \\ s_2 & s_3 & s_4 \end{vmatrix}.$$

Como ya sabemos por el párrafo anterior,

$$s_1 = \sigma_1 = -a,$$
$$s_2 = \sigma_1^2 - 2\sigma_2 = a^2 - 2b,$$
$$s_3 = \sigma_1^3 - 3\sigma_1\sigma_2 + 3\sigma_3 = -a^3 + 3ab - 3c.$$

Aplicando la fórmula de Newton, y teniendo en cuenta que $\sigma_4 = 0$, hallamos también que

$$s_4 = \sigma_1^4 - 4\sigma_1^2\sigma_2 + 4\sigma_1\sigma_3 + 2\sigma_2^2 = a^4 - 4a^2b + 4ac + 2b^2.$$

De aquí,

$$D = 3s_2s_4 + 2s_1s_2s_3 - s_2^3 - s_1^2s_4 - 3s_3^2 = a^2b^2 - 4b^3 - 4a^3c + 18abc - 27c^2. \tag{19}$$

En particular, siendo $a = 0$, o sea, para el polinomio cúbico incompleto, resulta:

$$D = -4b^3 - 27c^2,$$

lo cual está en correspondencia con lo que se dijo en el § 38.

§ 55. Segunda demostración del teorema fundamental del álgebra de los números complejos

La demostración del teorema fundamental, expuesta en el § 23, se efectuó de un modo no algebraico. Aquí queremos exponer otra demostración, en la que se emplea esencialmente el método algebraico. Así, pues, se aplicará el teorema fundamental de los polinomios simétricos (§ 52), y también el teorema de la existen-

cia de un campo de descomposición para cualquier polinomio (§ 48). Por otra parte, la parte no algebraica de la demostración será mínima y se reducirá a una afirmación muy sencilla.

Obsérvese primero que en el § 23 se demostró el lema del módulo del término superior de un polinomio. Suponiendo que los coeficientes del polinomio $f(x)$ son reales y poniendo $k = 1$, de este lema obtenemos el siguiente **corolario**:

Para valores reales de x suficientemente grandes en valor absoluto, el signo de un polinomio $f(x)$ de coeficientes reales coincide con el signo de su término superior.

De aquí se desprende el resultado siguiente:

Un polinomio de grado impar, de coeficientes reales, tiene por lo menos una raíz real.

En efecto, sea

$$f(x) = a_0x^n + a_1x^{n-1} + \ldots + a_n,$$

donde todos los coeficientes son reales. Como n es impar, el término superior a_0x^n, para valores positivos y negativos de x, tiene diferentes signos, por lo cual, como se ha demostrado más arriba, para valores positivos y negativos de x, suficientemente grandes en valor absoluto, el polinomio $f(x)$ también tiene signos distintos. Por consiguiente, existen unos valores reales de x, por ejemplo, a y b, tales que

$$f(a) < 0, \quad f(b) > 0.$$

Sin embargo, por el curso de análisis se sabe, que el polinomio $f(x)$ (o sea, la función racional entera) es una función continua y, por esto, en virtud de una de las principales propiedades de las funciones continuas, para ciertos valores reales de x comprendidos entre a y b, $f(x)$ toma cualquier valor previamente asignado, intermedio entre $f(a)$ y $f(b)$. En particular, existe un α, comprendido entre a y b, tal que $f(\alpha) = 0$.

Basándonos en este resultado, demostraremos ahora la proposición siguiente:

Todo polinomio de coeficientes reales, de un grado arbitrario, tiene por lo menos una raíz compleja.

En efecto, sea dado un polinomio $f(x)$ de coeficientes reales y de grado $n = 2^kq$, donde q es un número impar. Como el caso $k = 0$ ya se ha estudiado antes, supondremos que $k > 0$, o sea, que n es un número par, y haremos la demostración por inducción sobre k, suponiendo que nuestra afirmación ya está demostrada para **todos** los polinomios de coeficientes reales, cuyos grados son divisibles por 2^{k-1}, pero no son divisibles por 2^k*.

* Por consiguiente, estos grados pueden ser incluso mayores que n.

Sea P un campo de descomposición del polinomio $f(x)$ sobre el campo de los números complejos (véase el § 49) y sean $\alpha_1, \alpha_2, \ldots, \alpha_n$ las raíces de $f(x)$ contenidas en el campo P. Tomemos un número real arbitrario c y consideremos los elementos del campo P que son de la forma

$$\beta_{ij} = \alpha_i\alpha_j + c(\alpha_i + \alpha_j), \quad i < j. \tag{1}$$

Evidentemente, el número de elementos β_{ij} es igual a

$$\frac{n(n-1)}{2} = \frac{2^k q(2^k q - 1)}{2} = 2^{k-1}q(2^k q - 1) = 2^{k-1}q', \tag{2}$$

donde q' es un número impar.

Formemos ahora un polinomio $g(x)$ del anillo $P[x]$ que tenga por raíces todos estos elementos β_{ij} y sólo éstos:

$$g(x) = \prod_{i,\, j,\, i<j} (x - \beta_{ij}).$$

Los coeficientes de este polinomio son polinomios simétricos elementales en β_{ij}. Por consiguiente, en virtud de (1), son polinomios en $\alpha_1, \alpha_2, \ldots, \alpha_n$ de coeficientes reales (puesto que el número c es real) y, además, son simétricos. En efecto, la trasposición de cualesquiera dos α, por ejemplo, de α_k y α_l, implica solamente una permutación en el sistema de todas las β_{ij}; cualquiera β_{kj}, donde j es distinto de k y de l, se convierte en β_{lj} y viceversa, mientras que β_{kl} y todas las β_{ij}, para i y j diferentes de k y l, se quedan *en el sitio*. Mas, *los coeficientes del polinomio $g(x)$ no varían al permutar sus raíces.*

En virtud del teorema fundamental de los polinomios simétricos, de aquí se deduce que los coeficientes del polinomio $g(x)$ son polinomios (de coeficientes reales) en los coeficientes del polinomio dado $f(x)$ y, por esto, ellos mismos son números reales. El grado de este polinomio, igual al número de las raíces β_{ij}, en virtud de (2), es divisible por 2^{k-1}, pero no lo es por 2^k. Por esto, por la hipótesis de la inducción, al menos una de las raíces β_{ij} del *polinomio $g(x)$ tiene que ser un número complejo.*

Por lo tanto, cualquiera que sea el número real elegido c, se puede indicar un par de índices i, j, donde $1 \leqslant i \leqslant n$, $1 \leqslant j \leqslant n$, de modo que el elemento $\alpha_i\alpha_j + c(\alpha_i + \alpha_j)$ sea un número complejo; recordemos, que el campo P contiene al campo de los números complejos como subcampo. Se entiende que, por lo general, para otra elección del número c, a éste le va a corresponder en el sentido indicado otro par de índices. Sin embargo, existe una infinidad de números reales c distintos, mientras que nosotros *disponemos solamente de un número finito de pares i, j distintos.* De aquí se deduce, que se pueden elegir dos números reales **distin-**

tos c_1 y c_2, $c_1 \neq c_2$, tales, que a éstos les corresponde **un mismo par** de índices, para los cuales, los números

$$\left.\begin{aligned} \alpha_i\alpha_j + c_1(\alpha_i+\alpha_j) = a, \\ \alpha_i\alpha_j + c_2(\alpha_i+\alpha_j) = b \end{aligned}\right\} \qquad (3)$$

son complejos.

De la igualdad (3), resulta:

$$(c_1 - c_2)(\alpha_i + \alpha_j) = a - b,$$

de donde se deduce que

$$\alpha_i + \alpha_j = \frac{a-b}{c_1-c_2},$$

o sea, esta suma es un número complejo. De aquí, y si se quiere de la primera de las igualdades (3), se deduce que el producto $\alpha_i\alpha_j$ también es un número complejo. Por lo tanto, resulta que los elementos α_i y α_j son raíces de la ecuación cuadrática

$$x^2 - (\alpha_i + \alpha_j)x + \alpha_i\alpha_j = 0,$$

de coeficientes complejos, por lo cual, como esto se deduce de la fórmula para la resolución de la ecuación cuadrática con coeficientes complejos, obtenida en el § 38, ellos mismos tienen que ser números complejos. Por consiguiente, entre las raíces del polinomio $f(x)$ hemos hallado incluso dos complejas, con lo cual queda demostrada nuestra afirmación.

Para demostrar por completo el teorema fundamental, queda por considerar el caso de un polinomio de coeficientes complejos arbitrarios. Sea

$$f(x) = a_0x^n + a_1x^{n-1} + \ldots + a_n$$

un polinomio de este tipo. Consideremos el polinomio

$$\bar{f}(x) = \bar{a}_0x^n + \bar{a}_1x^{n-1} + \ldots + \bar{a}_n,$$

obtenido de $f(x)$ por sustitución de todos los coeficientes por sus conjugados, y examinemos el producto

$$F(x) = f(x)\bar{f}(x) = b_0x^{2n} + b_1x^{2n-1} + \ldots + b_kx^{2n-k} + \ldots + b_{2n},$$

donde, evidentemente,

$$b_k = \sum_{i+j=k} a_i\bar{a}_j, \quad k = 0, 1, 2, \ldots, 2n.$$

Basándose en las propiedades de los números complejos conjugados, conocidas por el § 18, obtenemos que

$$\bar{b}_k = \sum_{i+j=k} \bar{a}_i a_j = b_k,$$

o sea, todos los coeficientes del polinomio $F(x)$ son números reales.

Como se ha demostrado más arriba, de aquí se deduce que el polinomio $F(x)$ tiene por lo menos una raíz compleja β,

$$F(\beta)=f(\beta)\bar{f}(\beta)=0,$$

o sea, o $f(\beta)=0$, o bien, $\bar{f}(\beta)=0$. En el primer caso, el teorema queda demostrado. Si es que se cumple el segundo caso, o sea, si

$$\bar{a}_0\beta^n+\bar{a}_1\beta^{n-1}+\ldots+\bar{a}_n=0,$$

entonces, sustituyendo todos los números que figuran aquí por sus conjugados (que, como ya sabemos, no infringe la igualdad), obtenemos:

$$f(\bar{\beta})=a_0\bar{\beta}^n+a_1\bar{\beta}^{n-1}+\ldots+a_n=0,$$

o sea, el número complejo $\bar{\beta}$ es raíz de $f(x)$. La demostración del teorema fundamental se ha terminado.

CAPITULO XII

POLINOMIOS DE COEFICIENTES RACIONALES

§ 56. Reducibilidad de los polinomios sobre el campo de los números racionales

El tercer campo numérico que, junto con los campos de números reales y de números complejos tiene para nosotros un interés especial, *es el campo de los números racionales*; éste lo designaremos mediante R. Entre todos los campos numéricos éste es el más pequeño, pues, como se demostró en el § 43, el campo R está contenido totalmente en cualquier campo numérico. Ahora nos va a interesar el problema de la reducibilidad de los polinomios sobre el campo de *números racionales* y, *en el siguiente párrafo*, el problema de las raíces racionales (enteras o fraccionarias) de los polinomios de coeficientes racionales. Subrayemos una vez más, que éstos son dos problemas distintos; por ejemplo, el polinomio

$$x^4+2x^2+1=(x^2+1)^2$$

es reducible sobre el campo de números racionales, a pesar de que no tiene ninguna raíz racional.

¿Qué se puede decir de la reducibilidad de los polinomios sobre el campo R? Ante todo, obsérvese que, dado un polinomio $f(x)$ *de coeficientes racionales* que no sean *todos enteros, entonces*, reduciendo éstos a un común denominador y multiplicando $f(x)$ por este denominador, igual, por ejemplo, a k, resulta un polinomio $kf(x)$ cuyos coeficientes son ya números enteros. Es evidente, que los polinomios $f(x)$ y $kf(x)$ tienen raíces iguales; por otra parte, *éstos son a la vez reducibles o irreducibles sobre el campo* R.

Mas, por ahora, no tenemos derecho de limitarnos a estudiar en adelante los polinomios de coeficientes enteros. En efecto, supongamos que el polinomio $g(x)$ de coeficientes enteros es reducible sobre el campo de los números racionales, o sea, que se descompone *en factores de menor grado de coeficientes racionales (en general,* fraccionarios). ¿Se deduce de esto que $g(x)$ se descompone en factores de coeficientes enteros? En otras palabras, ¿puede ocurrir que un polinomio de coeficientes enteros sea reducible sobre el campo de números racionales y sea irreducible sobre el anillo de *los números enteros*?

La respuesta a estas preguntas se puede obtener haciendo un exámen análogo al que se hizo en el § 51. Llamemos *primitivo* al polinomio $f(x)$ de coeficientes enteros, si sus coeficientes son primos entre sí, o sea, si no tienen divisores comunes distintos de 1 y —1. Cualquier polinomio $\varphi(x)$ de coeficientes racionales se puede representar de un modo único en forma de un producto de una fracción irreducible por un polinomio primitivo:

$$\varphi(x) = \frac{a}{b} f(x); \tag{1}$$

para esto hay que sacar fuera de paréntesis el común denominador de todos los coeficientes del polinomio $\varphi(x)$, y después, los factores comunes de los numeradores de estos coeficientes; obsérvese que el grado de $f(x)$ es igual al grado de $\varphi(x)$. La unicidad (salvo el signo) de la representación (1) se demuestra del modo siguiente. Sea

$$\varphi(x) = \frac{a}{b} f(x) = \frac{c}{d} g(x),$$

donde $g(x)$ es de nuevo un polinomio primitivo. Entonces,

$$adf(x) = bcg(x).$$

Por lo tanto, ad y bc se han obtenido sacando todos los factores comunes de los coeficientes de un mismo polinomio de coeficientes enteros, por lo cual, pueden diferenciarse entre sí solamente en el signo. De aquí se deduce, que los polinomios primitivos $f(x)$ y $g(x)$ también pueden diferenciarse entre sí solamente en el signo.

Para los polinomios primitivos de coeficientes enteros conserva su valor el **lema de Gauss:**

El producto de dos polinomios primitivos de coeficientes enteros es un polinomio primitivo.

En efecto, sean dados los polinomios primitivos de coeficientes enteros

$$f(x) = a_0x^k + a_1x^{k-1} + \ldots + a_ix^{k-i} + \ldots + a_k,$$
$$g(x) = b_0x^l + b_1x^{l-1} + \ldots + b_jx^{l-j} + \ldots + b_l$$

y sea

$$f(x)\,g(x) = c_0x^{k+l} + c_1x^{k+l-1} + \ldots + c_{i+j}x^{(k+l)-(i+j)} + \ldots + c_{k+l}.$$

Si este producto no es primitivo, existe un número **primo** p *que* es común divisor de todos los coeficientes $c_0, c_1, \ldots, c_{k+l}$. Como no todos los coeficientes del polinomio primitivo $f(x)$ pueden dividirse por p, habrá uno, sea éste a_i, que será el primero que no se divide por p; del mismo modo, sea b_j el primer coeficiente del polinomio $g(x)$ que no se divide por p. Multiplicando término a término $f(x)$ por $g(x)$ y reuniendo los términos que contienen a $x^{(k+l)-(i+j)}$,

resulta:

$$c_{i+j} = a_i b_j + a_{i-1} b_{j+1} + a_{i-2} b_{j+2} + \ldots + a_{i+1} b_{j-1} + a_{i+2} b_{j-2} + \ldots$$

El primer miembro de esta igualdad se divide por p. Por éste se dividen también todos los términos del segundo miembro, menos el primero; en efecto, en virtud de las condiciones impuestas a la elección de i y j, todos los coeficientes $a_{i-1}, a_{i-2}, \ldots$, y también $b_{j-1}, b_{j-2}, \ldots$, se dividen por p. De esto se deduce, que el producto $a_i b_j$ también se divide por p y, por esto, como p es un número primo, tiene que dividirse por p por lo menos uno de los coeficientes a_i, b_j, lo cual, sin embargo, no es cierto. Con esto queda terminada la demostración del lema.

Pasemos a responder a las preguntas que se hicieron más arriba. Supongamos que el polinomio $g(x)$ de grado n, de coeficientes enteros, es reducible sobre el campo de números racionales:

$$g(x) = \varphi_1(x)\,\varphi_2(x),$$

donde $\varphi_1(x)$ y $\varphi_2(x)$ son polinomios de coeficientes racionales de grado menor que n. Entonces,

$$\varphi_i(x) = \frac{a_i}{b_i} f_i(x), \quad i = 1,\ 2,$$

donde $\frac{a_i}{b_i}$ es una fracción irreducible, $f_i(x)$ es un polinomio primitivo. Por lo tanto,

$$g(x) = \frac{a_1 a_2}{b_1 b_2} [f_1(x)\, f_2(x)].$$

El primer miembro de esta igualdad es un polinomio de coeficientes enteros, por esto, el denominador $b_1 b_2$ del segundo miembro tiene que simplificarse. Mas, por el lema de Gauss, el polinomio que figura entre corchetes es primitivo, por lo tanto, cualquier factor primo de $b_1 b_2$ puede simplificarse solamente con cierto factor primo de $a_1 a_2$, y como a_i y b_i son primos entre sí, $i = 1, 2$, el número a_2 tiene que dividirse por b_1 y el número a_1, por b_2:

$$a_2 = b_1 a_2', \quad a_1 = b_2 a_1'.$$

De aquí que

$$g(x) = a_1' a_2' f_1(x)\, f_2(x).$$

Uniendo el coeficiente $a_1' a_2'$ a cualquiera de los factores $f_1(x)$, $f_2(x)$, obtenemos la descomposición del polinomio $g(x)$ en factores de menor grado de coeficientes enteros. Con esto, queda demostrado el siguiente **teorema**:

Un polinomio de coeficientes enteros que es irreducible sobre el anillo de los números enteros, es irreducible también sobre el campo de los números racionales.

Por fin, hemos obtenido ahora el derecho de limitarnos a estudiar las descomposiciones de los polinomios de coeficientes enteros en factores cuyos coeficientes también sean enteros, en las cuestiones relacionadas con la irreducibilidad de los polinomios sobre el campo de números racionales.

Ya sabemos que sobre el campo de los números complejos, es reducible todo polinomio cuyo grado sea mayor que la unidad, y sobre el campo de los números reales, todo polinomio (de coeficientes reales) cuyo grado sea mayor que dos. Otra cosa ocurre en el caso del campo de los números racionales: *para cualquier n se puede indicar un polinomio de n-ésimo grado de coeficientes racionales (e incluso enteros) que es irreducible sobre el campo de los números racionales*. La demostración de esta afirmación se basa en el siguiente criterio suficiente de irreducibilidad de un polinomio sobre el campo R, denominado **criterio de Eisenstein:**

Sea dado un polinomio

$$f(x) = a_0x^n + a_1x^{n-1} + \ldots + a_{n-1}x + a_n,$$

de coeficientes enteros. Si, al menos de un modo, se puede elegir un número primo p que satisfaga a las condiciones siguientes:

1) *el coeficiente superior a_0 no es divisible por p,*

2) *todos los demás coeficientes son divisibles por p,*

3) *el término independiente, siendo divisible por p, no es divisible por p^2, entonces el polinomio $f(x)$ es irreducible sobre el campo de los números racionales.*

En efecto, si el polinomio $f(x)$ es reducible sobre el campo R, entonces se descompone en dos factores de menor grado de coeficientes enteros:

$$f(x) = (b_0x^k + b_1x^{k-1} + \ldots + b_k)(c_0x^l + c_1x^{l-1} + \ldots + c_l),$$

donde $k < n$, $l < n$, $k + l = n$. Identificando los coeficientes de ambos miembros de esta igualdad, obtenemos:

$$\left.\begin{aligned} a_n &= b_kc_l, \\ a_{n-1} &= b_kc_{l-1} + b_{k-1}c_l, \\ a_{n-2} &= b_kc_{l-2} + b_{k-1}c_{l-1} + b_{k-2}c_l, \\ &\cdots\cdots\cdots\cdots\cdots \\ a_0 &= b_0c_0. \end{aligned}\right\} \tag{2}$$

De la primera de las igualdades (2) se deduce que, como a_n es divisible por p y el número p es primo, uno de los factores b_k, c_l tiene que ser divisible por p. Ambos no pueden ser divisibles por p, puesto que, por la hipótesis, a_n no es divisible por p^2. Supongamos, por ejemplo, que b_k es divisible por p y, por lo tanto, c_l es primo

con p. Examinemos ahora la segunda de las igualdades (2). Su primer miembro, y también el primer término del segundo miembro, son divisibles por p, por lo cual, el producto $b_{k-1}c_l$ también es divisible por p; pero como c_l no es divisible por p, tiene que ser divisible por p el número b_{k-1}. De un modo semejante, de la tercera de las igualdades (2), resulta que b_{k-2} es divisible por p, etc. Por fin, de la $(k+1)$-ésima igualdad resultará que b_0 es divisible por p; pero entonces, de la última de las igualdades (2) se deduce que a_0 es divisible por p, lo cual contradice a la hipótesis.

Para cualquier n es muy fácil escribir polinomios de coeficientes enteros de n-ésimo grado que satisfagan a las condiciones del criterio de Eisenstein y, por lo tanto, que sean irreducibles sobre el campo de los números racionales. Tal es, por ejemplo, el polinomio $x^n + 2$; a éste es aplicable el criterio de Eisenstein para $p = 2$.

El criterio de Eisenstein es solamente una condición suficiente de irreducibilidad sobre el campo R, pero no es una condición necesaria: puede ocurrir que, para un polinomio dado $f(x)$, no se pueda elegir un número primo p, de modo que se cumplan las condiciones del criterio de Eisenstein, siendo el polinomio reducible como, por ejemplo, $x^2 - 5x + 6$, o irreducible, como $x^2 + 1$. Además del criterio de Eisenstein existen muchos más criterios suficientes distintos de irreducibilidad de los polinomios sobre el campo R que, por cierto, son menos importantes. Existe también un método que pertenece a Kronecker, que permite responder para cualquier polinomio de coeficientes enteros si éste es reducible o no lo es sobre el campo R. Mas, este método es muy complicado y casi no tiene aplicación práctica.

Ejemplo. Examinemos el polinomio

$$f_p(x) = \frac{x^p - 1}{x - 1} = x^{p-1} + x^{p-2} + \ldots + x + 1,$$

donde p es un número primo. Son raíces de este polinomio las raíces p-ésimas de la unidad, distintas de la unidad misma; como estas raíces, junto con la unidad, dividen al círculo unidad del campo complejo en p partes iguales, el polinomio $f_p(x)$ se llama *polinomio de división del círculo*.

A este polinomio no se le puede aplicar directamente el criterio de Eisenstein. Mas, hagamos una sustitución de la indeterminada, poniendo $x = y + 1$. Resulta:

$$g(y) = f_p(y+1) = \frac{(y+1)^p - 1}{(y-1) - 1} =$$

$$= \frac{1}{y}\left[y^p + py^{p-1} + \frac{p(p-1)}{2!}y^{p-2} + \ldots + py\right] =$$

$$= y^{p-1} + py^{p-2} + \frac{p(p-1)}{2!}y^{p-3} + \ldots + p.$$

Los coeficientes de polinomio $g(y)$ son los números binomiales y, por esto, todos, menos el superior, son divisibles por p; el término independiente no es divisible

por p^2. Por lo tanto, según el criterio de Eisenstein el polinomio $g(y)$ es irreducible sobre el campo R. De aquí se deduce la *irreducibilidad sobre el campo R del polinomio de división del círculo $f_p(x)$*. En efecto, si

$$f_p(x) = \varphi(x)\,\psi(x),$$

entonces

$$g(y) = \varphi(y+1)\,\psi(y+1).$$

§ 57. Raíces racionales de los polinomios de coeficientes enteros

Más arriba se señaló, que el problema de la descomposición de un polinomio dado en factores irreducibles sobre el campo de los números racionales no tiene prácticamente una solución más o menos satisfactoria. Pero un caso particular de este problema, referente a la separación de los factores lineales de un polinomio de coeficientes racionales, o sea, a la averiguación de sus raíces racionales, es muy elemental y se resuelve sin recurrir a cálculos complicados. Es comprensible que, con el problema de la averiguación de las raíces racionales de los polinomios de coeficientes racionales no se agota de ningún modo el problema general de las raíces reales de estos polinomios, es decir, que los métodos y resultados expuestos en el capítulo noveno conservan también enteramente su valor para los polinomios de coeficientes racionales.

Empezando a resolver el problema de la averiguación de las raíces racionales de los polinomios de coeficientes racionales, señalemos que, como se había indicado en el párrafo anterior, podemos limitarnos a estudiar solamente los polinomios de coeficientes enteros; además, se van a examinar por separado los casos de raíces enteras y de raíces fraccionarias.

Si el número entero α es raíz del polinomio $f(x)$ de coeficientes enteros, α es divisor del término independiente de este polinomio.

En efecto, sea

$$f(x) = a_0x^n + a_1x^{n-1} + \ldots + a_n.$$

Dividamos $f(x)$ por $x-\alpha$:

$$f(x) = (x-\alpha)(b_0x^{n-1} + b_1x^{n-2} + \ldots + b_{n-1}).$$

Efectuando la división por el método de Horner, expuesto en el § 22, obtenemos que *todos los coeficientes del cociente, incluyendo también b_{n-1}, son números enteros*, y como

$$a_n = -\alpha b_{n-1} = \alpha(-b_{n-1}),$$

nuestra proposición queda demostrada *.

* Sería erróneo demostrar este teorema alegando al hecho de que el término independiente a_n es el producto (salvo el signo) de todas las raíces del polinomio $f(x)$, pues, entre éstas puede haber también fraccionarias, irracionales y complejas, debido a lo cual, no se puede afirmar por anticipado que el producto de todas estas raíces, a excepción de α, es un número entero.

Por lo tanto, si un polinomio $f(x)$ de coeficientes enteros tiene raíces enteras, éstas se hallan entre los divisores del término independiente. Por consiguiente, se deben ensayar todos los divisores posibles del término independiente, tanto los positivos como los negativos; si ninguno de éstos es raíz del polinomio, este último carece en general de raíces.

Puede ocurrir que el ensayo de todos los divisores del término independiente sea muy engorroso, incluso cuando los valores del polinomio se calculen por el método de Horner en vez de sustituir directamente cada uno de los divisores en lugar de la indeterminada. Las observaciones que se hacen a continuación permiten simplificar un poco estos cálculos. Como 1 y -1 siempre son divisores del término independiente, se calculan en primer lugar $f(1)$ y $f(-1)$, lo cual no ofrece dificultad alguna. Si, luego, el número entero α es raíz de $f(x)$:

$$f(x)=(x-\alpha)\,q(x),$$

como se indicó más arriba, todos los coeficientes del cociente $q(x)$ son números enteros y, por esto, los cocientes

$$\frac{f(1)}{\alpha-1}=-q(1),\quad \frac{f(-1)}{\alpha+1}=-q(-1)$$

tienen que ser números enteros. Por lo tanto, *solamente tienen que ensayarse los divisores α del término independiente* (*distintos de 1 y -1*) *para los cuales cada uno de los cocientes $\frac{f(1)}{\alpha-1}$, $\frac{f(-1)}{\alpha+1}$ es un número entero.*

Ejmplos. 1. Hallar las raíces enteras del polinomio

$$f(x)=x^3-2x^2-x-6.$$

Los divisores del término independiente son: ±1, ±2, ±3, ±6. Como $f(1)=-8$, $f(-1)=-8$, los números 1 y -1 no son raíces. Por otra parte, los números

$$\frac{-8}{2+1},\ \frac{-8}{-2-1},\ \frac{-8}{6-1},\ \frac{-8}{-6-1}$$

son fraccionarios, por lo cual, los divisores 2, -2, 6, -6 tienen que ser desechados, mientras que los números

$$\frac{-8}{3-1},\ \frac{-8}{3+1},\ \frac{-8}{-3-1},\ \frac{-8}{-3+1}$$

son enteros, y por esto, los divisores 3 y -3 tienen que ser ensayados. Apliquemos el método de Horner:

$$-3\left|\begin{array}{cccc} 1 & -2 & -1 & -6 \\ \hline 1 & -5 & 14 & -48 \end{array}\right.,$$

o sea, $f(-3)=-48$ y, por esto, -3 no es raíz de $f(x)$. Finalmente,

$$3\left|\begin{array}{cccc} 1 & -2 & -1 & -6 \\ \hline 1 & 1 & 2 & 0 \end{array}\right.,$$

o sea, $f(3)=0$: el número 3 es raíz de $f(x)$. A la vez, hemos hallado los coeficientes del cociente de la división de $f(x)$ por $x-3$;

$$f(x)=(x-3)(x^2+x+2).$$

Fácilmente se observa que el número 3 no es raíz del cociente x^2+x+2, o sea, este número no es raíz múltiple de $f(x)$.

2. Hallar las raíces enteras del polinomio

$$f(x)=3x^4+x^3-5x^2-2x+2.$$

Aquí, los divisores del término independiente son: ± 1 y ± 2. Por otra parte, $f(1)=-1$, $f(-1)=1$, o sea, 1 y -1 no son raíces. Finalmente, como los números

$$\frac{1}{2+1} \text{ y } \frac{-1}{-2-1}$$

son fraccionarios, los números 2 y -2 tampoco serán raíces, por lo cual, el polinomio $f(x)$ carece de raíces enteras.

Examinemos el problema de las raíces fraccionarias.

Si un polinomio de coeficientes enteros, cuyo coeficiente superior es igual a la unidad, tiene una raíz racional, ésta es un número entero.

En efecto, supongamos que la fracción irreducible $\frac{b}{c}$ es raíz del polinomio

$$f(x)=x^n+a_1x^{n-1}+a_2x^{n-2}+\ldots+a_n$$

de coeficientes enteros, o sea, que

$$\frac{b^n}{c^n}+a_1\frac{b^{n-1}}{c^{n-1}}+a_2\frac{b^{n-2}}{c^{n-2}}+\ldots+a_n=0.$$

De aquí, resulta la igualdad

$$\frac{b^n}{c}=-a_1b^{n-1}-a_0b^{n-2}c-\ldots-a_nc^{n-1},$$

es decir, que una fracción irreducible es igual a un número entero, lo cual es imposible.

Para obtener todas las raíces racionales (enteras o fraccionarias) de un polinomio de coeficientes enteros

$$f(x)=a_0x^n+a_1x^{n-1}+a_2x^{n-2}+\ldots+a_{n-1}x+a_n$$

hay que hallar todas las raíces enteras del polinomio

$$\varphi(y)=y^n+a_1y^{n-1}+a_0a_2y^{n-2}+\ldots+a_0^{n-2}a_{n-1}y+a_0^{n-1}a_n$$

y dividirlas por a_0.

En efecto, multipliquemos $f(x)$ por a_0^{n-1}, y hagamos después la sustitución de la indeterminada poniendo $y=a_0x$. Evidentemente,

$$\varphi(y)=\varphi(a_0x)=a_0^{n-1}f(x).$$

De aquí se deduce, que las raíces del polinomio $f(x)$ son iguales a las raíces del polinomio $\varphi(y)$, divididas por a_0. En particular,

a las raíces racionales de $f(x)$ corresponderán raíces racionales de $\varphi(y)$; pero, como el coeficiente superior de $\varphi(y)$ es igual a la unidad, estas raíces sólo pueden ser enteras, y ya tenemos un método para buscarlas.

Ejemplo. Hallar las raíses racionales del polinomio

$$f(x)=3x^4+5x^3+x^2+5x-2.$$

Multiplicando $f(x)$ por 3^3 y poniendo $y=3x$, obtenemos:

$$\varphi(y)=y^4+5y^3+3y^2+45y-54.$$

Buscamos las raíces enteras del polinomio $\varphi(y)$.

Por el método de Horner, hallamos $\varphi(1)$:

$$1\left|\frac{1\ \ 5\ \ 3\ \ 45\ \ -54}{1\ \ 6\ \ 9\ \ 54\ \ \ \ \ 0}\right..$$

Por lo tanto, $\varphi(1)=0$, o sea, 1 es una raíz de $\varphi(y)$, siendo

$$\varphi(y)=(y-1)\,q(y),$$

donde

$$q(y)=y^3+6y^2+9y+54.$$

Hallemos las raíces enteras del polinomio $q(y)$. Los divisores del término independiente son: ± 1, ± 2, ± 3, ± 6, ± 9, ± 18, ± 27, ± 54. Aquí

$$q(1)=70, \quad q(-1)=50.$$

Calculando $\frac{q(1)}{\alpha-1}$ y $\frac{q(-1)}{\alpha+1}$ para cada divisor de α, se observa, que se tienen que desechar todos los divisores menos $\alpha=-6$. Ensayamos este divisor:

$$-6\left|\frac{1\ \ 6\ \ 9\ \ 54}{1\ \ 0\ \ 9\ \ \ 0}\right..$$

Por lo tanto, $q(-6)=0$, o sea, -6 es raíz de $q(y)$ y, por esto, de $\varphi(y)$.

Por consiguiente, el polinomio $\varphi(y)$ tiene las raíces enteras 1 y -6. Así, las raíces racionales del polinomio $f(x)$ son los números $\frac{1}{3}$ y -2, y sólo éstos.

Es menester subrayar una vez más, que los métodos expuestos anteriormente **solamente** se pueden aplicar a los polinomios de coeficientes enteros y **sólo** para hallar sus raíces racionales.

§ 58. Los números algebraicos

Todo polinomio de grado n de coeficientes racionales tiene n raíces en el campo de los números complejos, algunas de las cuales (e incluso todas) pueden estar fuera del campo de los números racionales. Mas, no cualquier número real o complejo es raíz de algún polinomio de coeficientes racionales. Los números complejos (y, en particular, los números reales) que son raíces de tales polinomios, se llaman números *algebraicos*, en contraposición a los números *trascendentes*. Entre los números algebraicos figuran los números

racionales, como raíces de los polinomios de primer grado de coeficientes racionales, y también cualquier radical de la forma $\sqrt[n]{a}$, siendo el subradical a un número racional, pues es raíz del binomio $x^n - a$. Por otra parte, en los cursos completos de análisis matemático se demuestra que es trascendente el número e, base del sistema de los logaritmos naturales, y también el número π, bien conocido en la geometría elemental.

Si el número α es algebraico, éste será incluso raíz de un polinomio de coeficientes enteros y, por esto, será raíz de uno de los divisores irreducibles de este polinomio, que también es de coeficientes enteros. *El polinomio irreducible de coeficientes enteros que tiene por raíz al número α se determina unívocamente, salvo un factor constante, o sea, de un modo único en absoluto, si se exige que los coeficientes de este polinomio sean primos entre sí* (es decir, que el polinomio sea primitivo). En efecto, si α es una raíz de dos polinomios irreducibles $f(x)$ y $g(x)$, el máximo común divisor de éstos tiene que ser distinto de la unidad, por lo cual, en virtud de su irreducibilidad, estos polinomios pueden diferenciarse entre sí solamente en un factor de grado cero.

Los números algebraicos que son raíces de un mismo polinomio irreducible (sobre el campo R), se llaman *conjugados entre sí**. Por consiguiente, todo el conjunto de números algebraicos se descompone en clases finitas disjuntas de números conjugados entre sí. Todo número racional, como raíz de un polinomio de primer grado, no tiene números conjugados distintos de sí mismo, siendo ésta una característica de los números racionales. En efecto, todo número algebraico que no sea racional será raíz de un polinomio irreducible de grado mayor que la unidad, por lo cual, tendrá algún conjugado distinto de sí mismo.

El conjunto de todos los números algebraicos es un subcampo del campo de los números complejos. En otras palabras, la suma, diferencia, producto y cociente de números algebraicos son también números algebraicos.

En efecto, supongamos que se han dado los números algebraicos α y β. Designemos mediante $\alpha_1 = \alpha, \alpha_2, \ldots, \alpha_n$, todos los números conjugados con α; mediante $\beta_1 = \beta, \beta_2, \ldots, \beta_s$, todos los números conjugados con β; mediante $f(x)$ y $g(x)$, los polinomios irreducibles de coeficientes racionales que tienen por raíces los números α y β, respectivamente. Escribamos un polinomio cuyas raíces sean todas las sumas posibles $\alpha_i + \beta_j$; éste es

$$\varphi(x) = \prod_{i=1}^{n} \prod_{j=1}^{s} [x - (\alpha_i + \beta_j)].$$

* No se debe confundir este concepto con el de números complejos conjugados.

Evidentemente, los coeficientes de este polinomio no varían al permutar entre sí todas las α_i, y también al permutar entre sí todas las β_j. Por consiguiente, según el teorema de los polinomios que son simétricos con respecto a dos sistemas de indeterminadas (véase el final del § 53), estos coeficientes son polinomios en los coeficientes de los polinomios $f(x)$ y $g(x)$. En otras palabras, resulta que los coeficientes del polinomio $\varphi(x)$ son números racionales, por lo cual, el número $\alpha+\beta=\alpha_1+\beta_1$, al ser una de sus raíces, es un número algebraico.

Del mismo modo, mediante los polinomios

$$\psi(x)=\prod_{i=1}^{n}\prod_{j=1}^{s}[x-(\alpha_i-\beta_j)]$$

y

$$\chi(x)=\prod_{i=1}^{n}\prod_{j=1}^{s}(x-\alpha_i\beta_j)$$

se demuestra que los números $\alpha-\beta$ y $\alpha\beta$ son algebraicos.

Para demostrar que el cociente de dos números algebraicos es un número algebraico, es suficiente demostrar que, si el número α es algebraico y distinto de cero, entonces, el número α^{-1} también lo es. Sea α raíz del polinomio

$$f(x)=a_0x^n+a_1x^{n-1}+\ldots+a_{n-1}x+a_n$$

de coeficientes racionales. Entonces, evidentemente, el polinomio

$$g(x)=a_nx^n+a_{n-1}x^{n-1}+\ldots+a_1x+a_0,$$

que también es de coeficientes racionales, tiene la raíz α^{-1}, como se quería demostrar.

Del teorema que acabamos de demostrar se deduce, que cualquier suma de un número racional y un radical, por ejemplo, $1+\sqrt[3]{2}$, y también cualquier suma de radicales, por ejemplo, $\sqrt{3}+\sqrt[7]{5}$, son números algebraicos. Mas, por ahora, no podemos afirmar que son algebraicos los números que se escriben en forma de radicales «de dos pisos», por ejemplo, $\sqrt{1+\sqrt{2}}$. Esto se va a deducir solamente del siguiente **teorema**:

Si el número ω es raíz del polinomio

$$\varphi(x)=x^n+\alpha x^{n-1}+\beta x^{n-2}+\ldots+\lambda x+\mu,$$

cuyos coeficientes son números algebraicos, entonces, ω es también un número algebraico.

Supongamos que α_i, β_j, ..., λ_s, μ_t toman todos los valores conjugados con los números α, β, ..., λ, μ, siendo $\alpha_1=\alpha$, $\beta_1=\beta$, ..., $\lambda_1=\lambda$, $\mu_1=\mu$. Consideremos todos los polinomios posibles

de la forma:

$$\varphi_{i, j, \ldots, s, t}(x) = x^n + \alpha_i x^{n-1} + \beta_j x^{n-2} + \ldots + \lambda_s x + \mu_t,$$

de modo que $\varphi_{1, 1, \ldots, 1, 1}(x) = \varphi(x)$, y tomemos el producto de todos estos polinomios:

$$F(x) = \prod_{i, j, \ldots, s, t} \varphi_{i, j, \ldots, s, t}(x).$$

Evidentemente, los coeficientes del polinomio $F(x)$ son simétricos con respecto a cada uno de los sistemas α_i, β_j, ..., λ_s, μ_t, por lo cual, (de nuevo en virtud del teorema del § 53), éstos son polinomios en los coeficientes de aquellos polinomios irreducibles (de *coeficientes racionales) cuyas raíces son* α, β, ..., λ, μ, *respectivamente*, o sea, ellos mismos son números racionales. Por consiguiente, el número ω, siendo raíz de $\varphi(x)$, es también raíz del polinomio $F(x)$ de coeficientes racionales, es decir, es un número algebraico.

Apliquemos este teorema al número $\omega = \sqrt{1 + \sqrt{2}}$. En virtud del teorema anterior, el número $\alpha = 1 + \sqrt{2}$ es algebraico y, por esto, el número ω es raíz del polinomio $x^2 - \alpha$, de coeficientes algebraicos, o sea, el mismo es algebraico. En general, reiterando los dos teoremas que acabamos de demostrar, el lector obtendrá sin dificultad alguna el siguiente resultado:

Todo número que se expresa por radicales sobre el campo de números racionales (es decir, que se expresa por una combinación de radicales, lo más complicada que sea, y en el caso general, por radicales «de muchos pisos»), es un número algebraico.

Evidentemente, los números algebraicos que se expresan por radicales forman un campo. Pero hay que tener presente que, como esto se deduce de la observación que se hizo (sin demostración) al final del § 38, éste es solamente una parte del campo de todos los números algebraicos.

Antes ya se había señalado que los números e y π son trascendentes. Pero, en la realidad, hay una infinidad de números trascendentes. Además, aplicando los conceptos y métodos de la teoría de los conjuntos, demostraremos que, en cierto sentido, hay más números trascendentes que algebraicos; el significado exacto de esta expresión quedará claro a continuación.

Un conjunto infinito M se llama *numerable*, si éste puede ponerse en correspondencia biunívoca con el conjunto de los números naturales, o sea, si sus elementos se pueden numerar mediante los números naturales, y *no numerable*, en caso contrario.

Lema 1. *Todo conjunto infinito M contiene un subconjunto numerable.*

En efecto, tomemos en M un elemento arbitrario a_1. Elijamos después un elemento a_2, distinto de a_1. En general, supongamos que ya se han elegido n elementos distintos en M: a_1, a_2, ... a_n. Como el conjunto M es infinito, éste no puede agotarse con los elementos elegidos, por lo cual, se puede indicar otro elemento a_{n+1}, distinto de éstos. Continuando este proceso, hallaremos en M un

subconjunto infinito formado por los elementos

$$a_1, a_2, \ldots, a_n, \ldots;$$

es evidente que este subconjunto es numerable.

Lema 2. *Todo subconjunto infinito B de un conjunto numerable A, es numerable.*

Como el conjunto A es numerable, éste se puede escribir en la forma:

$$a_1, a_2, \ldots, a_n, \ldots \tag{1}$$

Sea a_{k_1} el primer elemento de la sucesión (1) perteneciente a B; sea a_{k_2} el segundo elemento que tiene la misma propiedad, etc. Poniendo $a_{k_n} = b_n$, $n = 1, 2, \ldots$, obtenemos que los elementos del subconjunto B forman una sucesión,

$$b_1, b_2, \ldots, b_n, \ldots,$$

o sea, este subconjunto es numerable.

Lema 3. *La unión de un conjunto numerable de conjuntos finitos que no tienen elementos comunes, es un conjunto numerable.*

En efecto, sean dados los conjuntos finitos

$$A_1, A_2, \ldots, A_n, \ldots$$

y sea B la unión de ellos. Está claro que quedan numerados todos los elementos del conjunto B, si de un modo arbitrario se numeran los elementos del conjunto finito A_1, y después se continúa esta numeración pasando a considerar los elementos del conjunto A_2, etc.

Lema 4. *La unión de dos conjuntos numerables que no tienen elementos comunes, es un conjunto numerable.*

Sean dados los conjuntos numerables A con los elementos

$$a_1, a_2, \ldots, a_n, \ldots$$

y B con los elementos

$$b_1, b_2, \ldots, b_n, \ldots$$

y sea C la unión de estos conjuntos. Si se pone

$$a_n = c_{2n-1}, \; b_n = c_{2n}, \qquad n = 1, 2, \ldots,$$

todos los elementos del conjunto C quedarán representados en forma de la sucesión

$$c_1, c_2, \ldots, c_{2n-1}, c_{2n}, \ldots,$$

lo que demuestra que este conjunto es numerable.

Demostremos ahora el siguiente **teorema:**

El conjunto de todos los números algebraicos es numerable.

Demostremos previamente que *es numerable el conjunto de todos los polinomios en una indeterminada de coeficientes enteros.* Si

$$f(x) = a_0x^n + a_1x^{n-1} + \ldots + a_{n-1}x + a_n$$

es un polinomio de éstos, distinto de cero, llamaremos *altura* del polinomio al número natural

$$h_f = n + |a_0| + |a_1| + \ldots + |a_{n-1}| + |a_n|.$$

Es evidente, que existe solamente un número finito de polinomios de coeficientes enteros de una altura dada h; designemos este conjunto mediante M_h. Designemos también con M_0 el conjunto formado por el cero solamente. El conjunto de todos los polinomios de coeficientes enteros es la unión del conjunto numerable de los conjuntos finitos $M_0, M_1, M_2, \ldots, M_h, \ldots$, o sea, en virtud del lema 3, es numerable.

De aquí, por el lema 2, se deduce que *el conjunto de todos los polinomios primitivos irreducibles de coeficientes enteros también es numerable*. Por otra parte, ya sabemos que todo número algebraico es raíz de un polinomio primitivo irreducible de coeficientes enteros y solamente de uno. Por consiguiente, reuniendo las raíces de todos los polinomios de este tipo, o sea, tomando la unión de un conjunto numerable de conjuntos finitos, obtenemos el conjunto de todos los números algebraicos; por lo tanto, en virtud del lema 3, este conjunto es numerable.

Finalmente, demostremos el **teorema**:

El conjunto de todos los números trascendentes no es numerable.

Examinemos primero el conjunto F de todos los números reales x, situados entre el cero y la unidad, $0 < x < 1$, y demostremos que *este conjunto no es numerable*. Es sabido, que cada uno de los números indicados x se puede expresar en forma de una fracción decimal propia infinita

$$x = 0,\ \alpha_1\alpha_2 \ldots \alpha_n \ldots$$

y que esta expresión es única, si no se permiten fracciones en las que, para todos los n, empezando desde cierto $n = N$, todos los $\alpha_n = 9$; recíprocamente, cualquier fracción de la forma indicada es igual a cierto número x de este conjunto F. Supongamos ahora que el conjunto F es numerable, o sea, que todos los números x se pueden escribir en forma de una sucesión

$$x_1,\ x_2,\ \ldots,\ x_k,\ \ldots \tag{2}$$

Sea

$$x_k = 0,\ \alpha_{k_1}\alpha_{k_2} \ldots \alpha_{kn} \ldots$$

la expresión del número x_k en forma de fracción decimal infinita. Escribamos ahora una fracción decimal infinita

$$0,\ \beta_1\beta_2 \ldots \beta_n \ldots, \tag{3}$$

de modo que la cifra β_1 sea distinta de la primera cifra decimal de la fracción x_1, o sea, $\beta_1 \neq \alpha_{11}$, que la cifra β_2 sea distinta de la segunda cifra decimal de la fracción x_2, o sea, $\beta_2 \neq \alpha_{22}$, y, en general, que $\beta_n \neq \alpha_{nn}$. Supongamos además que entre las cifras β_n hay una infinidad de ellas, distintas de la cifra 9. Está claro que existe una fracción (3) que satisface a todas estas condiciones. Por consiguiente, ésta es un número del conjunto F y, por la construcción misma, es distinta de todos los números de la sucesión (2). Esta contradicción muestra que el conjunto F no es numerable.

De aquí se deduce, que *el conjunto de todos los números complejos no es numerable*, pues, en caso contrario, en virtud del lema 2, éste no podría contener el subconjunto no numerable F. En virtud del lema 4, es evidente ahora que no es numerable el conjunto de todos los números trascendentes, pues, la unión de este conjunto con el conjunto numerable de todos los números algebraicos es el conjunto de todos los números complejos, o sea, no es numerable.

En virtud del lema 1, los dos teoremas que hemos demostrado muestran que, en la realidad, el conjunto de los números trascendentes es más rico en elementos, o sea, es más «potente» que el conjunto de los números algebraicos.

CAPITULO XIII

FORMA NORMAL DE UNA MATRIZ

§ 59. Equivalencia de las λ-matrices

Aquí volvemos a examinar otra vez algunas cuestiones relacionadas con el álgebra lineal. Al estudiar el capítulo 7, el lector ya se habrá convencido del papel importante que desempeña el concepto de semejanza de las matrices. Precisando, dos matrices cuadradas de orden n son semejantes cuando, y sólo cuando, determinan (en bases diversas) una misma transformación lineal del espacio lineal de n dimensiones. Sin embargo, por ahora, no sabemos contestar a la pregunta, si son semejantes o no dos matrices determinadas. Por otra parte, no sabemos hallar, por ahora, entre todas las matrices semejantes a la matriz dada A, la que, en tal o cual sentido, tiene la forma más simple; incluso la cuestión sobre las condiciones para que una matriz A sea semejante a una matriz diagonal, fue estudiada en el § 33 solamente para un caso particular. Precisamente estas cuestiones se van a estudiar en el presente capítulo, y además, para el caso de un campo fundamental P arbitrario.

Ocupémonos primero del estudio de las matrices cuadradas de orden n, cuyos elementos son polinomios de grados arbitrarios en una indeterminada λ con coeficientes del campo P. Tales matrices se llaman *matrices polinomiales* o, abreviadamente, *λ-matrices*. Es un ejemplo de λ-matriz la matriz característica $A - \lambda E$ de una matriz cuadrada arbitraria A con elementos del campo P; en la diagonal principal de esta matriz figuran polinomios de primer grado; fuera de la diagonal principal, polinomios de grado cero o ceros. Cualquier matriz con elementos del campo P (para abreviar, a tales *matrices* las llamaremos *numéricas*) también será un caso particular de las λ-matrices: sus elementos son polinomios de grado cero, o son iguales a cero.

Sea dada una λ-matriz

$$A(\lambda) = \begin{pmatrix} a_{11}(\lambda) & \dots & a_{1n}(\lambda) \\ \dots & \dots & \dots \\ a_{n1}(\lambda) & \dots & a_{nn}(\lambda) \end{pmatrix}.$$

Llamemos *transformaciones elementales* de esta matriz a las transformaciones de los cuatro tipos siguientes:

1) multiplicación de cualquier fila de la matriz $A(\lambda)$ por cualquier número α del campo P, distinto de cero;

2) *multiplicación de cualquier columna* de la *matriz* $A(\lambda)$ *por* cualquier número α del campo P, distinto de cero;

3) agregación a cualquier i-ésima fila de la matriz $A(\lambda)$ una j-ésima fila cualquiera, $j \neq i$, y además, multiplicada por cualquier polinomio $\varphi(\lambda)$ del anillo $P[\lambda]$;

4) agregación a cualquier i-ésima columna de la matriz $A(\lambda)$ una j-ésima columna cualquiera, $j \neq i$, y además, multiplicada por cualquier polinomio $\varphi(\lambda)$ del anillo $P[\lambda]$.

Fácilmente se observa que, *para cada una de las transformaciones elementales de una* λ*-matriz, existe la transformación inversa, que también es elemental*. Así, pues, para la transformación 1), la inversa es la transformación elemental que consiste en multiplicar la misma fila por el número α^{-1}, que existe en virtud de la condición $\alpha \neq 0$; para la transformación 3), la inversa es la transformación que consiste en agregar a la i-ésima fila la j-ésima fila, multiplicada por $-\varphi(\lambda)$.

Efectuando unas cuantas transformaciones elementales en una matriz $A(\lambda)$, *se pueden permutar dos filas o dos columnas cualesquiera.*

Supongamos, por ejemplo, que se necesita permutar la i-ésima y la j-ésima filas de la matriz $A(\lambda)$. Como muestra el esquema que sigue, esto se realiza efectuando cuatro transformaciones elementales:

$$\begin{pmatrix} i \\ j \end{pmatrix} \to \begin{pmatrix} i+j \\ j \end{pmatrix} \to \begin{pmatrix} i+j \\ -i \end{pmatrix} \to \begin{pmatrix} j \\ -i \end{pmatrix} \to \begin{pmatrix} j \\ i \end{pmatrix}.$$

Aquí se ejecutaron las siguientes transformaciones: a) a la i-ésima fila se le agregó la j-ésima; b) de la j-ésima fila se restó la nueva i-ésima; c) a la nueva i-ésima fila se le agregó la nueva j-ésima; d) la nueva j-ésima fila se multiplicó por -1.

Diremos que las λ*-matrices* $A(\lambda)$ *y* $B(\lambda)$ *son equivalentes*, lo cual escribiremos con la notación $A(\lambda) \sim B(\lambda)$, si se puede pasar de la matriz $A(\lambda)$ a la matriz $B(\lambda)$ efectuando un número finito de transformaciones elementales. Es evidente que esta relación de equivalencia es reflexiva, transitiva y también simétrica, en virtud de la existencia de la transformación elemental inversa para cualquier transformación elemental. En otras palabras, *todas las* λ*-matrices cuadradas de orden* n *sobre el campo* P *se descomponen en clases disjuntas de matrices equivalentes.*

Nuestro objetivo próximo consiste en buscar, entre todas las λ-matrices equivalentes a una matriz dada $A(\lambda)$, una matriz que sea

lo más simple posible. Para esto, introduciremos el concepto siguiente. Se llama *λ-matriz canónica* a una λ-matriz que posea las tres propiedades siguientes:

a) esta matriz es diagonal, o sea, tiene la forma siguiente

$$\begin{pmatrix} e_1(\lambda) & & & & 0 \\ & e_2(\lambda) & & & \\ & & \cdot & & \\ & & & \cdot & \\ & & & & \cdot \\ 0 & & & & e_n(\lambda) \end{pmatrix} \tag{1}$$

b) cualquier polinomio $e_i(\lambda)$, $i = 2, 3, \ldots, n$, es divisible por el polinomio $e_{i-1}(\lambda)$;

c) el coeficiente superior de cada polinomio $e_i(\lambda)$, $i = 1, 2, \ldots$ $\ldots, n$, es igual a la unidad, si el polinomio es distinto de cero.

Obsérvese que, si entre los polinomios $e_i(\lambda)$ que figuran en la diagonal principal de la λ-matriz canónica (1), hay algunos iguales a cero, entonces, en virtud de la propiedad b), éstos inevitablemente ocupan los últimos sitios en la diagonal principal. Por otra parte, si entre los polinomios $e_i(\lambda)$ hay algunos de grado cero, entonces, según la propiedad c), éstos son todos iguales a 1 y, en virtud de la propiedad b), ocupan los primeros sitios en la diagonal principal de la matriz (1).

En particular, algunas matrices numéricas, como la matriz unidad y la matriz cero, son también λ-matrices canónicas.

Toda λ-matriz es equivalente a una λ-matriz canónica, o sea, en otras palabras, mediante transformaciones elementales se reduce a la forma canónica.

Demostraremos este teorema por inducción sobre el orden n de las λ-matrices consideradas. En efecto, para $n = 1$, se tiene:

$$A(\lambda) = (a(\lambda)).$$

Si $a(\lambda) = 0$, nuestra matriz ya es canónica. Si $a(\lambda) \neq 0$, es suficiente dividir el polinomio $a(\lambda)$ por su coeficiente superior —esto es una transformación elemental de la matriz— y obtenemos una matriz canónica.

Supongamos que el teorema ya está demostrado para las λ-matrices de orden $n - 1$. Examinemos una λ-matriz arbitraria $A(\lambda)$ de orden n. Si ésta es igual a cero, entonces ya es canónica y no hay nada que demostrar. Por esto, supondremos que entre los elementos de la matriz $A(\lambda)$ hay algunos distintos de cero.

Cambiando las filas de la matriz $A(\lambda)$ por columnas, si esto fuese necesario, se puede trasladar al ángulo superior de la izquierda uno de sus elementos distinto de cero. Por lo tanto, entre las λ-matri-

ces que son equivalentes a la matriz $A(\lambda)$, hay algunas en cuyos ángulos superiores de la izquierda figuran polinomios distintos de cero. Consideremos todas estas matrices. Los polinomios que figuran en el ángulo superior de la izquierda de estas matrices pueden tener grado distinto. Pero el grado de un polinomio es un número natural, y en cualquier conjunto de números naturales, no vacío, *existe el número menor. Por consiguiente, entre todas las λ-matrices* que son equivalentes a la matriz $A(\lambda)$ y que tienen en el ángulo superior de la izquierda un elemento distinto de cero, se puede hallar *una tal, que el polinomio que figure en dicho ángulo tenga el menor* grado posible. Finalmente, dividiendo la primera fila de esta matriz por el coeficiente superior del polinomio indicado, obtenemos una λ-matriz equivalente a la matriz $A(\lambda)$,

$$A(\lambda) \sim \begin{pmatrix} e_1(\lambda) & b_{12}(\lambda) & \dots & b_{1n}(\lambda) \\ b_{21}(\lambda) & b_{22}(\lambda) & \dots & b_{2n}(\lambda) \\ \dots & \dots & \dots & \dots \\ b_{n1}(\lambda) & b_{n2}(\lambda) & \dots & b_{nn}(\lambda) \end{pmatrix},$$

en la que $e_1(\lambda) \neq 0$, el coeficiente superior de este polinomio es igual a 1 y **con ninguna combinación de transformaciones elementales se puede pasar de la matriz obtenida a una matriz en cuyo ángulo superior de la izquierda figure un polinomio de grado menor, distinto de cero.**

Demostremos que *todos los elementos de la primera fila y de la primera columna de la matriz obtenida son divisibles por* $e_1(\lambda)$. Supongamos, por ejemplo, que, para $2 \leqslant j \leqslant n$,

$$b_{1j}(\lambda) = e_1(\lambda)\, q(\lambda) + r(\lambda),$$

donde el grado de $r(\lambda)$ es menor que el grado de $e_1(\lambda)$, si $r(\lambda)$ es diferente de cero. Entonces, restando de la j-ésima columna de nuestra matriz su primera columna, multiplicada por $q(\lambda)$, y permutando después la primera y j-ésima columnas, llegaremos a obtener una matriz equivalente a la matriz $A(\lambda)$, en cuyo ángulo superior de la izquierda figurará el polinomio $r(\lambda)$, o sea, un polinomio de grado menor que $e_1(\lambda)$, lo cual contradice a la elección de este polinomio. De aquí se deduce que $r(\lambda) = 0$, como se quería demostrar.

Restando ahora de la j-ésima columna de nuestra matriz su primera columna multiplicada por $q(\lambda)$, se sustituye el elemento $b_{1j}(\lambda)$ por cero. Realizando tales transformaciones para $j = 2, 3, \dots, n$, se sustituyen por ceros todos los elementos $b_{1j}(\lambda)$. De un modo análogo, se sustituyen también por ceros todos los elementos $b_{i1}(\lambda)$, $i = 2, 3, \dots, n$. *Por consiguiente, obtendremos una matriz equivalente $A(\lambda)$ en cuyo ángulo superior de la izquierda figurará*

el polinomio $e_1(\lambda)$ y en la que todos los demás elementos de la primera fila y de la primera columna serán iguales a cero:

$$A(\lambda) \sim \begin{pmatrix} e_1(\lambda) & 0 & \dots & 0 \\ 0 & c_{22}(\lambda) & \dots & c_{2n}(\lambda) \\ \dots & \dots & \dots & \dots \\ 0 & c_{n2}(\lambda) & \dots & c_{nn}(\lambda) \end{pmatrix}. \tag{2}$$

Por la hipótesis de inducción, la matriz de $(n-1)$-ésimo orden que figura en el ángulo inferior de la derecha de la matriz obtenida (2), mediante transformaciones elementales se reduce a la forma canónica:

$$\begin{pmatrix} c_{22}(\lambda) & \dots & c_{2n}(\lambda) \\ \dots & \dots & \dots \\ c_{n2}(\lambda) & \dots & c_{nn}(\lambda) \end{pmatrix} \sim \begin{pmatrix} e_2(\lambda) & & 0 \\ & \ddots & \\ 0 & & e_n(\lambda) \end{pmatrix}.$$

Efectuando las mismas transformaciones con las filas y columnas correspondientes de la matriz (2) —evidentemente, en este caso la primera fila y la primera columna de esta matriz se quedan invariables —, obtenemos que

$$A(\lambda) \sim \begin{pmatrix} e_1(\lambda) & & & 0 \\ & e_2(\lambda) & & \\ & & \ddots & \\ 0 & & & e_n(\lambda) \end{pmatrix}. \tag{3}$$

Para demostrar que la matriz (3) es canónica no queda más que demostrar que $e_2(\lambda)$ es divisible por $e_1(\lambda)$. Supongamos que

$$e_2(\lambda) = e_1(\lambda)\, q(\lambda) + r(\lambda),$$

donde $r(\lambda) \neq 0$ y el grado de $r(\lambda)$ es menor que el de $e_1(\lambda)$. Pero, agregando a la segunda columna de la matriz (3) su primera columna multiplicada por $q(\lambda)$ y restando después de la segunda fila la primera, *se sustituye el elemento $e_2(\lambda)$ por el elemento $r(\lambda)$*. Permutando luego las primeras dos filas y las primeras dos columnas, conseguiremos trasladar el polinomio $r(\lambda)$ al ángulo superior de la izquierda de la matriz, lo cual, sin embargo, contradice a la elección del polinomio $e_1(\lambda)$.

El teorema de la reducción de una λ-matriz a la forma canónica queda demostrado. Este teorema se puede completar con el siguiente **teorema de unicidad:**

Toda λ-matriz es equivalente solamente a una matriz canónica.

En efecto, sea dada una λ-matriz arbitraria $A(\lambda)$ de orden n. Fijemos algún número natural k, $1 \leqslant k \leqslant n$, y consideremos todos los menores de k-ésimo orden de la matriz $A(\lambda)$. Calculando estos menores obtenemos un sistema finito de polinomios en λ; designemos con $d_k(\lambda)$ el máximo común divisor de este sistema de polinomios, tomado con el coeficiente superior 1.

Por consiguiente, tenemos los polinomios

$$d_1(\lambda),\ d_2(\lambda),\ \ldots,\ d_n(\lambda), \tag{4}$$

determinados unívocamente por la misma matriz $A(\lambda)$. Aquí, $d_1(\lambda)$ es el máximo común divisor de todos los elementos de la matriz $A(\lambda)$, tomado con el coeficiente superior 1, y $d_n(\lambda)$ es igual al determinante de la matriz $A(\lambda)$, dividido por su coeficiente superior. Obsérvese también, que si la matriz $A(\lambda)$ tiene rango r, entonces

$$d_{r+1}(\lambda) = \ldots = d_n(\lambda) = 0,$$

mientras que todos los demás polinomios del sistema (4) son distintos de cero.

El máximo común divisor $d_k(\lambda)$ de todos los menores de k-ésimo orden de una λ-matriz $A(\lambda)$, $k = 1, 2, \ldots, n$, no varía al realizar transformaciones elementales en la matriz $A(\lambda)$.

Esta proposición es casi evidente, si se efectúan transformaciones elementales del tipo 1) y 2) en la matriz $A(\lambda)$. Así, por ejemplo, si la i-ésima fila de la matriz se multiplica por un número α del campo P, $\alpha \neq 0$, todos los menores de k-ésimo orden, por los que pasa la i-ésima fila, se multiplicarán por α, mientras que los demás menores de k-ésimo orden se quedarán invariables. Mas, al buscar el máximo común divisor de unos cuantos polinomios, cualesquiera de éstos se pueden multiplicar por números del campo P distintos de cero.

Examinemos ahora las transformaciones elementales del tipo 3) y 4). Supongamos, por ejemplo, que a la i-ésima fila de la matriz $A(\lambda)$ se le agrega su j-ésima fila, $j \neq i$, multiplicada por el polinomio $\varphi(\lambda)$; designemos con $\overline{A}(\lambda)$ la matriz que resulta después de esta transformación y con $\overline{d}_k(\lambda)$, el máximo común divisor de todos sus menores de k-ésimo orden, tomado con el coeficiente superior 1. Veamos lo que ocurre con los menores de k-ésimo orden de la matriz $A(\lambda)$ al hacer esta transformación.

Está claro que no varían los menores por los que no pasa la i-ésima fila. Tampoco varían los menores por los que pasan la i-ésima y la j-ésima filas, pues el determinante no varía al sumar a una de sus filas un múltiplo de otra fila. Por fin, tomemos cualquiera de los menores de k-ésimo orden por los que pasa la i-ésima

fila, pero no pasa la j-ésima; designémoslo mediante M. Evidentemente, el menor correspondiente de la matriz $\overline{A}(\lambda)$ se puede representar en forma de una suma del menor M y de un menor M', multiplicado por $\varphi(\lambda)$, donde este último es el menor de la matriz $A(\lambda)$ que se obtiene del menor M al sustituir los elementos de la i-ésima fila de la matriz $A(\lambda)$ por sus elementos correspondientes de la j-ésima fila. Como M y M' son divisibles por $d_k(\lambda)$, también será divisible por $d_k(\lambda)$ la suma $M+\varphi(\lambda)M'$.

De lo dicho se deduce, que todos los menores de k-ésimo orden de la matriz $\overline{A}(\lambda)$ son divisibles por $d_k(\lambda)$, por lo cual, $\overline{d}_k(\lambda)$ también es divisible por $d_k(\lambda)$. Pero, como para la transformación elemental considerada existe una transformación elemental inversa del mismo tipo, $d_k(\lambda)$ también es disvisible por $\overline{d}_k(\lambda)$. Si se tiene en cuenta que los coeficientes superiores de estos polinomios son iguales a 1, se tiene $\overline{d}_k(\lambda)=d_k(\lambda)$, como se quería demostrar.

Por lo tanto, *a todas las λ-matrices equivalentes a la matriz $A(\lambda)$ corresponde una misma colección de polinomios* (4). En particular, esto mismo se refiere a cualquier (si hay varias) matriz canónica equivalente a $A(\lambda)$. Supongamos que (3) es una de estas matrices.

Calculemos el polinomio $d_k(\lambda)$, $k=1, 2, \ldots, n$, utilizando la matriz (3). Está claro, que el menor de k-ésimo orden que figura en el ángulo superior de la izquierda de esta matriz, es igual al producto

$$e_1(\lambda)\,e_2(\lambda)\,\ldots\,e_k(\lambda). \tag{5}$$

Si, luego, se toma en la matriz (3) el menor de k-ésimo orden que figura en las filas cuyos índices son $i_1, i_2, \ldots, i_k$, donde $i_1 < i_2 < \ldots < i_k$, y en las columnas *que tienen los mismos índices de ordenación*, resulta que este menor es igual al producto $e_{i_1}(\lambda)\,e_{i_2}(\lambda)\,\ldots\,e_{i_k}(\lambda)$, el cual es divisible por (5). En efecto, $1\leqslant i_1$, y, por esto, $e_{i_1}(\lambda)$ es divisible por $e_1(\lambda)$; $2\leqslant i_2$, y por esto, $e_{i_2}(\lambda)$ es divisible por $e_2(\lambda)$, etc. Finalmente, si en la matriz (3) *se toma el menor de k-ésimo orden por el que pasa, al menos para una i*, la i-ésima fila de esta matriz, pero no pasa su i-ésima columna, resulta que este menor contiene una fila nula, por lo cual, es igual a cero.

De lo expuesto se deduce, que el producto (5) es el máximo común divisor de todos los menores de k-ésimo orden de la matriz (3) y, por consiguiente, de la matriz inicial $A(\lambda)$,

$$d_k(\lambda)=e_1(\lambda)\,e_2(\lambda)\,\ldots\,e_k(\lambda), \qquad k=1, 2, \ldots, n. \tag{6}$$

Ahora es fácil demostrar que *los polinomios $e_k(\lambda)$, $k=1, 2, \ldots, n$, se determinan unívocamente por la misma matriz $A(\lambda)$*. **Supongamos que el rango de esta matriz es** r. Entonces, como ya sabemos,

$d_r(\lambda) \neq 0$, pero $d_{r+1}(\lambda) = 0$, y por esto, en virtud de (6), $e_{r+1}(\lambda) = 0$. De aquí, en virtud de las propiedades de la matriz canónica, se deduce, en general, que **si el rango r de la matriz $A(\lambda)$ es menor que n, entonces,**

$$e_{r+1}(\lambda) = e_{r+2}(\lambda) = \ldots = e_n(\lambda) = 0. \tag{7}$$

Por otra parte, para $k \leqslant r$, como $d_{k-1}(\lambda) \neq 0$, de (6) resulta que

$$e_k(\lambda) = \frac{d_k(\lambda)}{d_{k-1}(\lambda)}. \tag{8}$$

Con esto se termina la demostración de la unicidad de la forma canónica de una λ-matriz.

Al mismo tiempo hemos obtenido un método para hallar directamente los polinomios $e_k(\lambda)$ llamados *factores invariantes* de la matriz $A(\lambda)$.

Ejemplo. Reducir a la forma canónica la λ-matriz

$$A(\lambda) = \begin{pmatrix} \lambda^3 - \lambda & 2\lambda^2 \\ \lambda^2 + 5\lambda & 3\lambda \end{pmatrix}.$$

Efectuando una cadena de transformaciones elementales, obtenemos:

$$A(\lambda) \sim \begin{pmatrix} \lambda^3 - \lambda & \frac{2}{3}\lambda^2 \\ \lambda^2 + 5\lambda & \lambda \end{pmatrix} \sim \begin{pmatrix} \frac{1}{3}\lambda^3 - \frac{10}{3}\lambda^2 - \lambda & 0 \\ \lambda^2 + 5\lambda & \lambda \end{pmatrix} \sim$$

$$\sim \begin{pmatrix} \frac{1}{3}\lambda^3 - \frac{10}{3}\lambda^2 - \lambda & 0 \\ 0 & \lambda \end{pmatrix} \sim \begin{pmatrix} \lambda^3 - 10\lambda^2 - 3\lambda & 0 \\ 0 & \lambda \end{pmatrix} \sim$$

$$\sim \begin{pmatrix} \lambda & 0 \\ 0 & \lambda^3 - 10\lambda^2 - 3\lambda \end{pmatrix}.$$

Pero se podrían calcular directamente los factores invariantes de la matriz $A(\lambda)$. Precisamente, calculando el máximo común divisor de los elementos de esta matriz, obtenemos:

$$d_1(\lambda) = e_1(\lambda) = \lambda.$$

Calculando el determinante de la matriz $A(\lambda)$ y observando que su coeficiente superior es igual a 1, resulta:

$$d_2(\lambda) = \lambda^4 - 10\lambda^3 - 3\lambda^2,$$

y, por esto,

$$e_2(\lambda) = \frac{d_2(\lambda)}{d_1(\lambda)} = \lambda^3 - 10\lambda^2 - 3\lambda.$$

§ 60. λ-matrices unimodulares. Relación entre la semejanza de las matrices numéricas y la equivalencia de sus matrices características

De los resultados del párrafo precedente se desprende un criterio de equivalencia de las λ-matrices, que se puede formular de los siguientes dos modos, que son casi idénticos:

Dos λ-matrices son equivalentes si, y sólo si, éstas se reducen a una misma forma canónica.

Dos λ-matrices son equivalentes si, y sólo si, éstas tienen factores invariantes iguales.

Deduzcamos otro criterio de carácter distinto.

Ya sabemos que al conjunto de las λ-matrices canónicas pertenece la matriz unidad E. Llamemos a una λ-matriz $U(\lambda)$ *unimodular*, si su forma canónica coincide con la matriz unidad E, o sea, si todos sus factores invariantes son iguales a la unidad.

Una λ-matriz $U(\lambda)$ es unimodular si, y sólo si, su determinante es distinto de cero, pero no depende de λ, o sea, si es un número del campo fundamental P, distinto de cero.

En efecto, si $U(\lambda) \sim E$, a estas dos matrices les corresponde un mismo polinomio $d_n(\lambda)$. Pero, para la matriz unidad, $d_n(\lambda) = 1$. De aquí se deduce que el determinante de la matriz $U(\lambda)$, que se diferencia de $d_n(\lambda)$ solamente en un factor numérico distinto de cero, es un número del campo P, distinto de cero. Recíprocamente, si el determinante de la matriz $U(\lambda)$ es diferente de cero y no depende de λ, entonces, para esta matriz, el polinomio $d_n(\lambda)$ será igual a 1, por lo cual, según (6) del párrafo anterior, todos los factores invariantes $e_i(\lambda)$ de la matriz $U(\lambda)$, $i = 1, 2, \ldots, n$, son iguales a la unidad.

De aquí se deduce que, *toda matriz numérica no degenerada es una λ-matriz unimodular*. Pero, una λ-matriz unimodular puede ser de forma complicada. Así, pues, la λ-matriz

$$\begin{pmatrix} \lambda & \lambda^3 + 5 \\ \lambda^2 - \lambda - 4 & \lambda^4 - \lambda^3 - 4\lambda^2 + 5\lambda - 5 \end{pmatrix}$$

es unimodular, pues su determinante es igual a 20, o sea, es distinto de cero y no depende de λ.

Del teorema demostrado anteriormente se deduce que, *el producto de λ-matrices unimodulares es unimodular*, pues, es suficiente recordar que, al multiplicar matrices, sus determinantes se multiplican.

Una λ-matriz $U(\lambda)$ es unimodular si, y sólo si, existe la matriz inversa y ésta es una λ-matriz.

En efecto, dada una λ-matriz no degenerada, buscando de un modo ordinario la matriz inversa, tendremos que dividir los complementos algebraicos de los elementos de la matriz dada por el determinante de ésta, o sea, por un polinomio en λ. Por esto, en el caso general, los elementos de la matriz inversa serán fracciones racionales en λ, pero no polinomios en λ, o sea, **esta matriz no será una λ-matriz.** Si se da una matriz unimodular, habrá que dividir los complementos algebraicos solamente por un número del campo P, distinto de cero, o sea, los elementos de la matriz inversa serán polinomios en λ, por lo cual, la misma matriz inversa será una λ-matriz. Recíproca-

mente, si una λ-matriz $U(\lambda)$ tiene λ-matriz inversa $U^{-1}(\lambda)$, los determinantes de ambas matrices serán polinomios en λ, su producto será igual a 1, por lo cual, ambos determinantes tendrán que ser polinomios de grado cero.

De la última observación, se deduce el siguiente complemento del teorema que acabamos de demostrar:

Una λ-matriz que es inversa a una λ-matriz unimodular, es también unimodular.

El concepto de matriz unimodular se emplea en el enunciado siguiente del nuevo **criterio de equivalencia de las λ-matrices:**

Dos λ-matrices $A(\lambda)$ y $B(\lambda)$ de orden n son equivalentes si, y sólo si, existen unas λ-matrices unimodulares $U(\lambda)$ y $V(\lambda)$ del mismo orden n, tales que

$$B(\lambda) = U(\lambda)\,A(\lambda)\,V(\lambda). \tag{1}$$

Introduzcamos primero el siguiente concepto, que se emplea en la demostración de este criterio. Llamemos *matriz elemental* a la matriz numérica (que, por lo tanto, es una λ-matriz):

$$\begin{pmatrix} 1 & & & & & & 0 \\ & \ddots & & & & & \\ & & 1 & & & & \\ \cdots & \cdots & \cdots & \alpha & \cdots & \cdots & \cdots \\ & & & & 1 & & \\ & & & & & \ddots & \\ 0 & & & & & & 1 \end{pmatrix} (i), \tag{2}$$

que se diferencia de la matriz unidad solamente en que, en cierto i-ésimo lugar de la diagonal principal, $1 \leqslant i \leqslant n$, figura un número arbitrario α del campo P, **distinto de cero.** Por otra parte, llamemos **también** *matriz elemental* a la λ-matriz

$$\begin{pmatrix} 1 & & & \vdots & & & \\ & \ddots & & \vdots & & & \\ \cdots & \cdots & 1 & \cdots & \varphi(\lambda) & \cdots & \cdots \\ & & & \vdots & & & \\ & & & 1 & & & \\ & & & \vdots & \ddots & & \\ & & & \vdots & & & 1 \end{pmatrix} (i), \tag{3}$$

(j)

que se diferencia de la matriz unidad solamente en que, en la intersección de la i-ésima fila y la j-ésima columna, $1 \leqslant i \leqslant n$, $1 \leqslant j \leqslant n$, siendo $i \neq j$, figura un polinomio arbitrario $\varphi(\lambda)$ del anillo $P[\lambda]$.

Toda matriz elemental es unimodular. En efecto, el determinante de la matriz (2) es igual a α, pero, por la condición, $\alpha \neq 0$; por otra parte, el determinante de la matriz (3) es igual a 1.

La ejecución en una λ-matriz $A(\lambda)$ de cualquier transformación elemental es equivalente a la multiplicación de esta matriz a la izquierda o a la derecha por una matriz elemental.

En efecto, el lector comprobará sin dificultad la justeza de las cuatro proposiciones siguientes: 1) multiplicar la matriz $A(\lambda)$ a la izquierda por la matriz (2) equivale a multiplicar la i-ésima de la matriz $A(\lambda)$ por el número α; 2) multiplicar la matriz $A(\lambda)$ a la derecha por la matriz (2) equivale a multiplicar la i-ésima columna de la matriz (2) por el número α; 3) multiplicar la matriz $A(\lambda)$ a la izquierda por la matriz (3) equivale a sumar a la i-ésima fila de la matriz $A(\lambda)$ su j-ésima fila, multiplicada por $\varphi(\lambda)$; 4) multiplicar la matriz $A(\lambda)$ a la derecha por la matriz (3) equivale a sumar a la j-ésima columna de la matriz $A(\lambda)$ su i-ésima columna, multiplicada por $\varphi(\lambda)$.

Pasemos a demostrar ahora nuestro **criterio de equivalencia de las λ-matrices.** Si $A(\lambda) \sim B(\lambda)$, de la matriz $A(\lambda)$ se puede pasar a la matriz $B(\lambda)$ realizando un número finito de transformaciones elementales. Sustituyendo cada una de estas transformaciones por la multiplicación a la izquierda o a la derecha, por una matriz elemental, llegaremos a la siguiente igualdad:

$$B(\lambda) = U_1(\lambda) \ldots U_k(\lambda) A(\lambda) V_1(\lambda) \ldots V_l(\lambda), \tag{4}$$

donde todas las matrices $U_1(\lambda), \ldots, U_k(\lambda), V_1(\lambda) \ldots, V_l(\lambda)$ son elementales y, por consiguiente, unimodulares. Por esto, serán también unimodulares las matrices

$$U(\lambda) = U_1(\lambda) \ldots U_k(\lambda), \quad V(\lambda) = V_1(\lambda) \ldots V_l(\lambda), \tag{5}$$

que son productos de matrices unimodulares, y la igualdad (4) se escribirá de la forma (1). Obsérvese que si, por ejemplo, $k = 0$, o sea, que se efectuaron transformaciones elementales solamente sobre las columnas, entonces, ponemos simplemente $U(\lambda) = E$.

La parte ya demostrada permite a la vez enunciar la siguiente proposición:

Una λ-matriz es unimodular si, y sólo si, ésta se representa en forma de un producto de matrices elementales.

En efecto, ya hemos empleado el hecho de que el producto de matrices elementales es unimodular. Recíprocamente, una matriz unimodular arbitraria $W(\lambda)$ es equivalente a la matriz unidad E.

Aplicando a las matrices E y $W(\lambda)$ la demostración que se llevó a cabo con las matrices $A(\lambda)$ y $B(\lambda)$, de (4) obtenemos la igualdad

$$W(\lambda)=U_1(\lambda)\ \dots\ U_k(\lambda)\,V_1(\lambda)\ \dots\ V_l(\lambda),$$

o sea, la matriz $W(\lambda)$ ha quedado representada en forma de un producto de matrices elementales.

Ahora es fácil demostrar **la proposición recíproca de nuestro criterio.** Supongamos que para las matrices $A(\lambda)$ y $B(\lambda)$ existen unas matrices unimodulares $U(\lambda)$ y $V(\lambda)$ tales, que se verifica la igualdad (1). Por lo demostrado, las matrices $U(\lambda)$ y $V(\lambda)$ se pueden representar en forma de productos de matrices elementales; supongamos que (5) son las representaciones dichas. La igualdad (1) se escribirá ahora en la forma (4) y, sustituyendo cada multiplicación por una matriz elemental por su transformación elemental correspondiente, obtenemos, por fin, que $A(\lambda)\sim B(\lambda)$.

Polinomios matriciales. El concepto de λ-matriz se puede interpretar de otro modo. Llamemos *λ-polinomio matricial de orden n sobre el campo P* a un polinomio en λ cuyos coeficientes son matrices cuadradas de un mismo orden n, con elementos del mismo campo P; su forma general es:

$$A_0\lambda^k+A_1\lambda^{k-1}+\dots+A_{k-1}\lambda+A_k. \tag{6}$$

Entendiendo el producto de la matriz A_i por λ^{k-i}, $i=0,1,\dots,k$, en correspondencia con el § 15, como el producto de todos los elementos de la matriz A_i por λ^{k-i}, y efectuando después la suma de las matrices de acuerdo con el mismo § 15, obtenemos que, *todo λ-polinomio matricial de orden n se puede expresar en forma de una λ-matriz de orden n*. Así, pues,

$$\begin{pmatrix}4&0\\-1&1\end{pmatrix}\lambda^3+\begin{pmatrix}0&-3\\0&1\end{pmatrix}\lambda^2+\begin{pmatrix}1&2\\0&-2\end{pmatrix}\lambda+\begin{pmatrix}0&1\\0&0\end{pmatrix}=$$
$$=\begin{pmatrix}4\lambda^3+\lambda&-3\lambda^2+2\lambda+1\\-\lambda^3&\lambda^3+\lambda^2-2\lambda\end{pmatrix}.$$

Recíprocamente, *toda λ-matriz de orden n se puede expresar en forma de un λ-polinomio matricial de orden n*. Así, pues,

$$\begin{pmatrix}3\lambda^2-5&\lambda+1\\\lambda^4+2\lambda&-3\end{pmatrix}=\begin{pmatrix}0&0\\1&0\end{pmatrix}\lambda^4+\begin{pmatrix}3&0\\0&0\end{pmatrix}\lambda^2+\begin{pmatrix}0&1\\2&0\end{pmatrix}\lambda+\begin{pmatrix}-5&1\\0&-3\end{pmatrix}.$$

La correspondencia entre las λ-matrices y los λ-polinomios matriciales es biunívoca e isomorfa en el sentido del § 46. En efecto, la igualdad de los λ-polinomios de la forma (6) como matrices es equivalente a la igualdad de los coeficientes matriciales de potencias iguales de λ, y la multiplicación de una matriz por λ es equivalente

a su multiplicación por una matriz escalar con λ en la diagonal principal.

Sea dada una λ-matriz $A(\lambda)$, siendo

$$A(\lambda) = A_0\lambda^k + A_1\lambda^{k-1} + \ldots + A_{k-1}\lambda + A_k,$$

donde la matriz A_0 no es nula. Al número k lo llamaremos *grado* de la λ-matriz $A(\lambda)$; evidentemente, éste será el grado superior (respecto a λ) de los elementos de la matriz $A(\lambda)$.

La consideración de las λ-matrices como polinomios matriciales permite desarrollar para las λ-matrices una teoría de divisibilidad análoga a la teoría de divisibilidad de los polinomios numéricos, pero, naturalmente, más complicada porque el producto de las matrices no es conmutativo y por la existencia de divisores de cero. Nos limitaremos a estudiar el **algoritmo de la división con resto.**

Sean dadas sobre el campo P las λ-matrices de orden n:

$$A(\lambda) = A_0\lambda^k + A_1\lambda^{k-1} + \ldots + A_{k-1}\lambda + A_k,$$

$$B(\lambda) = B_0\lambda^l + B_1\lambda^{l-1} + \ldots + B_{l-1}\lambda + B_l;$$

supongamos que la matriz B_0 no es degenerada, o sea, que existe la matriz B_0^{-1}. Entonces, sobre el campo P se pueden hallar unas λ-matrices $Q_1(\lambda)$ y $R_1(\lambda)$ del mismo orden n, tales que

$$A(\lambda) = B(\lambda)\,Q_1(\lambda) + R_1(\lambda), \qquad \textbf{(7)}$$

donde el grado de $R_1(\lambda)$ es menor que el grado de $B(\lambda)$, o bien, $R_1(\lambda) = 0$. Por otra parte, sobre el campo P se pueden hallar unas λ-matrices $Q_2(\lambda)$ y $R_2(\lambda)$ de orden n, tales que

$$A(\lambda) = Q_2(\lambda)\,B(\lambda) + R_2(\lambda), \qquad (8)$$

donde el grado de $R_2(\lambda)$ es menor que el grado de $B(\lambda)$, o bien, $R_2(\lambda) = 0$. Las matrices $Q_1(\lambda)$ y $R_1(\lambda)$, y también $Q_2(\lambda)$ y $R_2(\lambda)$, que satisfacen a estas condiciones, se determinan unívocamente.

La demostración de este teorema se efectúa del mismo modo que la demostración del teorema correspondiente para los polinomios numéricos (véase el § 20). Supongamos, por ejemplo, que a la condición (7) satisfacen también las matrices $\overline{Q}_1(\lambda)$ y $\overline{R}_1(\lambda)$, donde el grado de $\overline{R}_1(\lambda)$ es menor que el de $B(\lambda)$. Entonces,

$$B(\lambda)\,[Q_1(\lambda) - \overline{Q}_1(\lambda)] = \overline{R}_1(\lambda) - R_1(\lambda).$$

El grado del segundo miembro es menor que l, mientras que el grado del primer miembro es mayor o igual a l, si la expresión entre corchetes es diferente de cero, puesto que la matriz B_0 no es degenerada. De aquí se deduce la unicidad de las matrices $Q_1(\lambda)$ y $R_1(\lambda)$.

Para demostrar la existencia de estas matrices, observemos que, para $k \geqslant l$, el grado de la diferencia

$$A(\lambda) - B(\lambda) \cdot B_0^{-1} A_0 \lambda^{k-l}$$

es estrictamente menor que k; por esto, $B_0^{-1} A_0 \lambda^{k-l}$ será el término superior del λ-polinomio matricial $Q_1(\lambda)$. A continuación se obra igual que en el § 20. Por otra parte, el grado de la diferencia

$$A(\lambda) - A_0 B_0^{-1} \lambda^{k-l} \cdot B(\lambda)$$

también es estrictamente menor que k, o sea, $A_0 B_0^{-1} \lambda^{k-l}$ es el término superior del λ-polinomio matricial $Q_2(\lambda)$. Vemos, pues, que en el caso general, las λ-matrices $Q_1(\lambda)$ y $Q_2(\lambda)$ (y también $R_1(\lambda)$ y $R_2(\lambda)$), que satisfacen a las condiciones del teorema, verdaderamente, son distintas.

Teorema fundamental de la semejanza de las matrices. Como ya se señaló, todavía no conocemos un procedimiento para responder a la pregunta si unas matrices numéricas dadas A y B (o sea, matrices con elementos del campo fundamental P) son semejantes o no. Por otra parte, sus matrices características $A - \lambda E$ y $B - \lambda E$ son λ-matrices y el problema de la equivalencia de estas matrices se resuelve de un modo efectivo. Por esto, se comprende el valor tan grande que tiene el siguiente teorema:

Las matrices A y B, con elementos del campo P, son semejantes si, y sólo si, sus matrices características $A - \lambda E$ y $B - \lambda E$ son equivalentes.

En efecto, supongamos que las matrices A y B son semejantes, o sea, que sobre el campo P existe una matriz no degenerada C tal, que

$$B = C^{-1} A C.$$

Entonces

$$C^{-1}(A - \lambda E) C = C^{-1} A C - \lambda (C^{-1} E C) = B - \lambda E.$$

Pero las matrices numéricas no degeneradas C^{-1} y C son λ-matrices unimodulares. Vemos, pues, que la matriz $B - \lambda E$ se obtiene multiplicando la matriz $A - \lambda E$ a la izquierda y a la derecha por matrices unimodulares, o sea, $A - \lambda E \sim B - \lambda E$.

La demostración del teorema recíproco es más complicada. Supongamos que

$$A - \lambda E \sim B - \lambda E.$$

Entonces, existen unas matrices unimodulares $U(\lambda)$ y $V(\lambda)$, tales, que

$$U(\lambda)(A - \lambda E) V(\lambda) = B - \lambda E. \tag{9}$$

Teniendo en cuenta que para las matrices unimodulares existen las matrices inversas y éstas son λ-matrices, de (9) deducimos las

siguientes igualdades, que se emplean a continuación:

$$\left.\begin{aligned} U(\lambda)(A-\lambda E) &= (B-\lambda E)V^{-1}(\lambda), \\ (A-\lambda E)V(\lambda) &= U^{-1}(\lambda)(B-\lambda E). \end{aligned}\right\} \qquad (10)$$

Como la λ-matriz $B - \lambda E$ es de grado 1 con respecto a λ, y además, el coeficiente superior del polinomio matricial correspondiente es la matriz no degenerada $-E$, a las matrices $U(\lambda)$ y $B - \lambda E$ se les puede aplicar el algoritmo de la división con resto, según el cual, existen unas matrices $Q_1(\lambda)$ y R_1 (esta última, si es distinta de cero, tiene que ser de grado 0 con respecto a λ, o sea, no depende de λ), tales, que

$$U(\lambda) = (B-\lambda E)Q_1(\lambda) + R_1. \qquad (11)$$

De modo análogo,

$$V(\lambda) = Q_2(\lambda)(B-\lambda E) + R_2. \qquad (12)$$

Aplicando (11) y (12), de (9) obtenemos:

$$R_1(A-\lambda E)R_2 = (B-\lambda E) - U(\lambda)(A-\lambda E)Q_2(\lambda)(B-\lambda E) -$$
$$-(B-\lambda E)Q_1(\lambda)(A-\lambda E)V(\lambda) + (B-\lambda E)Q_1(\lambda)(A-\lambda E)Q_2(B-\lambda E)$$

o, en virtud de (10),

$$R_1(A-\lambda E)R_2 = (B-\lambda E) - (B-\lambda E)V^{-1}(\lambda)Q_2(\lambda)(B-\lambda E) -$$
$$-(B-\lambda E)Q_1(\lambda)U^{-1}(\lambda)(B-\lambda E) +$$
$$+(B-\lambda E)Q_1(\lambda)(A-\lambda E)Q_2(\lambda)(B-\lambda E) = (B-\lambda E)\times$$
$$\times\{E - [V^{-1}(\lambda)Q_2(\lambda) + Q_1(\lambda)U^{-1}(\lambda) - Q_1(\lambda)(A-\lambda E)Q_2(\lambda)](B-\lambda E)\}$$

La expresión que figura entre corchetes en el segundo miembro, verdaderamente, es igual a cero. En caso contrario, ésta, siendo una λ-matriz, puesto que $V^{-1}(\lambda)$, así como $U^{-1}(\lambda)$, son λ-matrices, sería por lo menos de grado cero y, entonces, el grado de la expresión entre llaves sería no menor que 1 y, por consiguiente, el grado de todo el segundo miembro no sería menor que 2. Pero, esto es imposible, puesto que en el primer miembro figura una λ-matriz de grado 1.

Por lo tanto,

$$R_1(A-\lambda E)R_2 = B - \lambda E,$$

de donde, igualando los coeficientes matriciales de potencias iguales de λ, obtenemos:

$$R_1AR_2 = B, \qquad (13)$$

$$R_1R_2 = E. \qquad (14)$$

La igualdad (14) muestra que la matriz numérica R_2 no sólo es distinta de cero, sino incluso no degenerada, siendo

$$R_2^{-1} = R_1,$$

y entonces, la igualdad (13) toma la forma

$$R_2^{-1} A R_2 = B,$$

lo cual demuestra la semejanza de las matrices A y B.

A la vez, **hemos aprendido a hallar la matriz R_2 no degenerada que transforma la matriz A en la matriz B.** Precisando, si las matrices $A - \lambda E$ y $B - \lambda E$ son equivalentes, entonces con un número finito de transformaciones elementales se transforma la primera en la segunda. Tomemos las transformaciones de éstas que se relacionan a las columnas, y designemos mediante $V(\lambda)$ el producto de las matrices elementales correspondientes, tomadas en el mismo orden. Dividamos después $V(\lambda)$ por $B - \lambda E$, de modo que el cociente quede a la izquierda del divisor (véase (8)). El residuo de esta división será la matriz R_2.

En realidad, se puede prescindir de la división indicada, utilizando el siguiente lema que hallará también aplicación en el § 62:

Lema. *Sea*

$$V(\lambda) = V_0\lambda^s + V_1\lambda^{s-1} + \ldots + V_{s-1}\lambda + V_s, \quad V_0 \neq 0. \tag{15}$$

Si

$$V(\lambda) = (\lambda E - B)\, Q_1(\lambda) + R_1, \tag{16}$$
$$V(\lambda) = Q_2(\lambda)\,(\lambda E - B) + R_2,$$

se tiene

$$R_1 = B^s V_0 + B^{s-1} V_1 + \ldots + B V_{s-1} + V_s, \tag{17}$$
$$R_2 = V_0 B^s + V_1 B^{s-1} + \ldots + V_{s-1} B + V_s.$$

Es suficiente demostrar, por ejemplo, la primera de estas dos afirmaciones, *pues la segunda se demuestra por analogía. La demostración* consiste en la comprobación directa del cumplimiento de la igualdad (16); para esto el polinomio $V(\lambda)$ se sustituye por su expresión (15), *en lugar de R_1 se pone (17) y en vez de $Q_1(\lambda)$ se toma* el polinomio

$$Q_1(\lambda) = V_0\lambda^{s-1} + (BV_0 + V_1)\lambda^{s-2} + (B^2V_0 + BV_1 + V_2)\lambda^{s-3} + \ldots$$
$$\ldots + (B^{s-1}V_0 + B^{s-2}V_1 + \ldots + V_{s-1}).$$

La prueba de esto la dejamos a cuenta del lector.

Ejemplo. Sean dadas las matrices

$$A=\begin{pmatrix}-2 & 1\\ 0 & 3\end{pmatrix},\quad B=\begin{pmatrix}-10 & -4\\ 26 & 11\end{pmatrix}.$$

Sus matrices características son equivalentes, puesto que se reducen a una misma forma canónica

$$\begin{pmatrix}1 & 0\\ 0 & \lambda^2-\lambda-6\end{pmatrix},$$

por esto, las matrices A y B son semejantes.

Para hallar la matriz R_2 que transforma A en B, hallemos alguna cadena de transformaciones elementales que transforme $A-\lambda E$ en $B-\lambda E$. Así, pues,

$$A-\lambda E=\begin{pmatrix}-2-\lambda & 1\\ 0 & 3-\lambda\end{pmatrix}\sim\begin{pmatrix}-2-\lambda & 1\\ -16-8\lambda & 11-\lambda\end{pmatrix}\sim$$

$$\sim\begin{pmatrix}8+4\lambda & -4\\ -16-8\lambda & 11-\lambda\end{pmatrix}\sim\begin{pmatrix}40+4\lambda & -4\\ -104 & 11-\lambda\end{pmatrix}\sim\begin{pmatrix}-10-\lambda & -4\\ 26 & 11-\lambda\end{pmatrix}=B-\lambda E.$$

Las dos últimas transformaciones se refieren a las columnas: a la primera columna se suma la segunda, multiplicada por -8, y después, la primera columna se multiplica por $-\frac{1}{4}$. El producto de las matrices elementales correspondientes es

$$V(\lambda)=\begin{pmatrix}1 & 0\\ -8 & 1\end{pmatrix}\begin{pmatrix}-\frac{1}{4} & 0\\ 0 & 1\end{pmatrix}=\begin{pmatrix}-\frac{1}{4} & 0\\ 2 & 1\end{pmatrix}.$$

Esta matriz no depende de λ, por lo cual, ésta será la matriz R_2 buscada.

Claro, la matriz que transforma A en B está muy lejos de determinarse unívocamente. Tal es también, por ejemplo, la matriz

$$\begin{pmatrix}3 & 1\\ 2 & 1\end{pmatrix}.$$

§ 61. Forma normal de Jordan

Ahora estudiaremos las matrices cuadradas de orden n con elementos del campo P. Se distinguirá un tipo especial de matrices de éstas, llamadas **matrices de Jordan**, y se demostrará que estas matrices sirven de forma normal para una clase de matrices muy amplia. Precisando, *las matrices cuyas raíces características pertenecen al campo fundamental P (y solamente tales matrices), son semejantes a ciertas matrices de Jordan, o, como suele decirse, se reducen a la forma normal de Jordan.* De aquí se deducirá que, si se toma por P el campo de los números complejos, *cualquier matriz con elementos complejos se reduce a la forma normal de Jordan en este campo.*

Introduzcamos las definiciones necesarias. Se llama *«malla» de Jordan* de orden k correspondiente al número λ_0, la matriz de orden

k, $1 \leqslant k \leqslant n$, que tiene la forma

$$\begin{pmatrix} \lambda_0 & 1 & & & & 0 \\ & \lambda_0 & 1 & & & \\ & & \ddots & \ddots & & \\ & & & \ddots & \ddots & \\ & & & & \ddots & 1 \\ 0 & & & & & \lambda_0 \end{pmatrix}; \qquad (1)$$

en otras palabras, en su diagonal principal figura un mismo número λ_0 del campo P; la paralela, situada por encima de la diagonal principal y más próxima a ésta, está ocupada totalmente por el número 1; todos los demás elementos de la matriz son iguales a cero. Así, pues,

$$(\lambda_0), \quad \begin{pmatrix} \lambda_0 & 1 \\ 0 & \lambda_0 \end{pmatrix}, \quad \begin{pmatrix} \lambda_0 & 1 & 0 \\ 0 & \lambda_0 & 1 \\ 0 & 0 & \lambda_0 \end{pmatrix}$$

son mallas de Jordan de primero, segundo y tercer orden, respectivamente.

Se llama *matriz de Jordan* de orden n a la matriz de order n que tiene la forma

$$J = \begin{pmatrix} \boxed{J_1} & & & 0 \\ & \boxed{J_2} & & \\ & & \ddots & \\ 0 & & & \boxed{J_s} \end{pmatrix}; \qquad (2)$$

aquí, a lo largo de la diagonal principal figuran las mallas de Jordan $J_1, J_2, \ldots, J_s$ de ciertos órdenes, que necesariamente no son distintas, y que corresponden a ciertos números del campo P, tampoco necesariamente distintos; todos los lugares fuera de estas mallas están ocupados por ceros. Aquí, $s \geqslant 1$, o sea, una malla de Jordan de orden n pertenece al conjunto de las matrices de Jordan de este mismo orden y, naturalmente, $s \leqslant n$.

Obsérvese, a pesar de que esto no se va a aplicar a continuación, que se podría haber descrito la estructura de la matriz de Jordan sin recurrir al concepto de célula de Jordan. Así, pues, es evidente que una matriz es de Jordan si, y sólo si, ésta tiene la forma

$$\begin{pmatrix} \lambda_1 & \varepsilon_1 & & & & 0 \\ & \lambda_2 & \varepsilon_2 & & & \\ & & \ddots & \ddots & & \\ & & & \ddots & \ddots & \\ & & & & \ddots & \varepsilon_{n-1} \\ 0 & & & & & \lambda_n \end{pmatrix},$$

donde λ_i, $i = 1, 2, \ldots, n$, son números arbitrarios del campo P, y cada ε_j, $j = 1, 2, \ldots, n-1$, es igual a la unidad o a cero, siendo $\lambda_j = \lambda_{j+1}$ cuando $\varepsilon_j = 1$.

Las matrices diagonales son casos particulares de las matrices de Jordan; éstas son precisamente las matrices de Jordan cuyas mallas son de orden 1.

Nuestro objetivo próximo consiste en **hallar la forma canónica para la matriz característica $J - \lambda E$ de una matriz arbitraria de Jordan J, de orden n.** Hallemos primero la forma canónica para la matriz característica

$$\begin{pmatrix} \lambda_0-\lambda & 1 & & & & 0 \\ & \lambda_0-\lambda & 1 & & & \\ & & \ddots & \ddots & & \\ & & & \ddots & \ddots & \\ & & & & \ddots & 1 \\ 0 & & & & & \lambda_0-\lambda \end{pmatrix} \tag{3}$$

de una malla de Jordan (1), de orden k. Calculando el determinante de esta matriz y recordando que el coeficiente superior del polinomio $d_k(\lambda)$ tiene que ser igual a 1, obtenemos,

$$d_k(\lambda) = (\lambda - \lambda_0)^k.$$

Por otra parte, entre los menores de $(k-1)$-ésimo orden de la matriz (3) hay uno que es igual a la unidad, precisamente, el que resulta después de haber suprimido la primera columna y la última fila de esta matriz. Por esto,

$$d_{k-1}(\lambda) = 1.$$

De aquí se deduce que *la forma canónica de la matriz* (3) *es la siguiente* λ*-matriz de orden* k:

$$\begin{pmatrix} 1 & & & 0 \\ & \cdot & & \\ & & \cdot & \\ & & & 1 \\ 0 & & & (\lambda-\lambda_0)^k \end{pmatrix}. \tag{4}$$

Demostremos ahora el **lema** que sigue:

Si los polinomios $\varphi_1(\lambda)$, $\varphi_2(\lambda)$, $\ldots$, $\varphi_t(\lambda)$ *del anillo* $P[\lambda]$ *son primos entre sí dos a dos, se verifica la siguiente equivalencia*:

$$\begin{pmatrix} \varphi_1(\lambda) & & & & 0 \\ & \varphi_2(\lambda) & & & \\ & & \cdot & & \\ & & & \cdot & \\ 0 & & & & \varphi_t(\lambda) \end{pmatrix} \sim \begin{pmatrix} 1 & & & & 0 \\ & \cdot & & & \\ & & \cdot & & \\ & & & 1 & \\ 0 & & & & \prod_{i=1}^{t} \varphi_i(\lambda) \end{pmatrix}.$$

Evidentemente, es suficiente considerar el caso $t = 2$. Como los polinomios $\varphi_1(\lambda)$ y $\varphi_2(\lambda)$ son primos entre sí, en el anillo $P[\lambda]$ existen unos polinomios $u_1(\lambda)$ y $u_2(\lambda)$, tales que

$$\varphi_1(\lambda)u_1(\lambda) + \varphi_2(\lambda)u_2(\lambda) = 1.$$

Por esto,

$$\begin{pmatrix} \varphi_1(\lambda) & 0 \\ 0 & \varphi_2(\lambda) \end{pmatrix} \sim \begin{pmatrix} \varphi_1(\lambda) & \varphi_1(\lambda)u_1(\lambda) \\ 0 & \varphi_2(\lambda) \end{pmatrix} \sim$$
$$\sim \begin{pmatrix} \varphi_1(\lambda) & \varphi_1(\lambda)u_1(\lambda)+\varphi_2(\lambda)u_2(\lambda) \\ 0 & \varphi_2(\lambda) \end{pmatrix} = \begin{pmatrix} \varphi_1(\lambda) & 1 \\ 0 & \varphi_2(\lambda) \end{pmatrix} \sim$$
$$\sim \begin{pmatrix} 1 & \varphi_1(\lambda) \\ \varphi_2(\lambda) & 0 \end{pmatrix} \sim \begin{pmatrix} 1 & \varphi_1(\lambda) \\ 0 & -\varphi_1(\lambda)\varphi_2(\lambda) \end{pmatrix} \sim$$
$$\sim \begin{pmatrix} 1 & 0 \\ 0 & -\varphi_1(\lambda)\varphi_2(\lambda) \end{pmatrix} \sim \begin{pmatrix} 1 & 0 \\ 0 & \varphi_1(\lambda)\varphi_2(\lambda) \end{pmatrix},$$

como se quería demostrar.

Consideremos ahora la matriz característica

$$J-\lambda E = \begin{pmatrix} \boxed{J_1-\lambda E_1} & & & & & 0 \\ & \boxed{J_2-\lambda E_2} & & & & \\ & & \cdot & & & \\ & & & \cdot & & \\ & & & & \cdot & \\ 0 & & & & & \boxed{J_s-\lambda E_s} \end{pmatrix} \tag{5}$$

de la matriz de Jordan J de la forma (2); aquí, E_i, $i = 1, 2, \ldots, s$, es la matriz unidad del mismo orden que la malla J_i. Supongamos que las mallas de Jordan de la matriz J corresponden a los siguientes números distintos: $\lambda_1, \lambda_2, \ldots, \lambda_t$, donde $t \leqslant s$. Supongamos también que al número λ_i, $i = 1, 2, \ldots, t$, corresponden q_i mallas de Jordan, $q_i \geqslant 1$, y sean los órdenes de éstas, colocados en orden no creciente, los números

$$k_{i1} \geqslant k_{i2} \geqslant \ldots \geqslant k_{iq_i}. \tag{6}$$

Obsérvese, a pesar de que no vamos a utilizar esto, que

$$\sum_{i=1}^{t} q_i = s,$$

$$\sum_{i=1}^{t} \sum_{j=1}^{q_i} k_{ij} = n.$$

Aplicando transformaciones elementales a las filas y columnas de la matriz (5) que pasan por la malla $J_i - \lambda E_i$ de esta matriz, no se tocan, evidentemente, las otras mallas diagonales. De aquí se deduce que en la matriz (5), mediante transformaciones elementales, se puede sustituir cada malla $J_i - \lambda E_i$, $i = 1, 2, \ldots, s$, por la malla correspondiente de la forma (4). En otras palabras, *la matriz $J - \lambda E$ es equivalente a una matriz diagonal, en cuya diagonal, además de ciertas unidades, figuran también los siguientes polinomios, que corresponden a todas las mallas de Jordan de la matriz J;*

$$\left.\begin{array}{l} (\lambda-\lambda_1)^{k_{11}},\ (\lambda-\lambda_1)^{k_{12}},\ \ldots,\ (\lambda-\lambda_1)^{k_{1q_1}}, \\ (\lambda-\lambda_2)^{k_{21}},\ (\lambda-\lambda_2)^{k_{22}},\ \ldots,\ (\lambda-\lambda_2)^{k_{2q_2}}, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ (\lambda-\lambda_t)^{k_{t1}},\ (\lambda-\lambda_t)^{k_{t2}},\ \ldots,\ (\lambda-\lambda_t)^{k_{tq_t}}. \end{array}\right\} \tag{7}$$

En este caso, no indicamos los lugares en la diagonal donde figuran los polinomios (7), puesto que en cualquier λ-matriz diagonal, los elementos diagonales se pueden cambiar de sitio arbitrariamente permutando las filas y las columnas homólogas. Esta observación se debe tener en cuenta a continuación.

Sea q el máximo entre los números q_i, $i = 1, 2, \ldots, t$. Designemos con $e_{n-j+1}(\lambda)$ el producto de los polinomios que figuran en la j-ésima columna de la tabla (7), $j = 1, 2, \ldots, q$, o sea,

$$e_{n-j+1}(\lambda) = \prod_{i=1}^{t} (\lambda - \lambda_i)^{k_{ij}}; \tag{8}$$

si en este caso en la j-ésima columna hay sitios vacíos (para algunos i puede ocurrir que $q_i < j$), suponemos que los factores correspon-

dientes en (8) son iguales a la unidad. Como, por la condición, los números $\lambda_1, \lambda_2, \ldots, \lambda_t$ son distintos, los grados de los binomios lineales que figuran en la j-ésima columna de la tabla (7) son primos entre sí dos a dos. Por esto, según el lema demostrado anteriormente, mediante transformaciones elementales, estos binomios se pueden sustituir en la matríz diagonal considerada por su producto $e_{n-j+1}(\lambda)$ y cierta cantidad de unidades. Haciendo esto para $j = 1, 2, \ldots, q$, obtenemos que

$$J-\lambda E \sim \begin{pmatrix} 1 & & & & & & & 0 \\ & \cdot & & & & & & \\ & & \cdot & & & & & \\ & & & 1 & & & & \\ & & & & e_{n-q+1}(\lambda) & & & \\ & & & & & \cdot & & \\ & & & & & & e_{n-1}(\lambda) & \\ 0 & & & & & & & e_n(\lambda) \end{pmatrix}. \tag{9}$$

Esta es la forma canónica buscada de la matriz $J-\lambda E$. En efecto, los coeficientes superiores de todos los polinomios que figuran en (9) en la diagonal principal, son iguales a la unidad, y, en virtud de la condición (6), cada uno de estos polinomios es divisible por el precedente.

Ejemplo. Sea

$$J=\begin{pmatrix} \boxed{\begin{matrix} 2 & 1 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 2 \end{matrix}} & & & & 0 \\ & \boxed{2} & & & \\ & & \boxed{\begin{matrix} 5 & 1 \\ 0 & 5 \end{matrix}} & & \\ & & & \boxed{\begin{matrix} 5 & 1 \\ 0 & 5 \end{matrix}} & \\ 0 & & & & \boxed{2} \end{pmatrix}$$

Para esta matriz de Jordan de noveno orden, la tabla de polinomios (7) es de la forma

$$(\lambda-2)^3,\ \lambda-2,\ \lambda-2,$$
$$(\lambda-5)^2,\ (\lambda-5)^2.$$

Por esto, los factores invariantes de la matriz J son:

$$e_9(\lambda) = (\lambda - 2)^3 (\lambda - 5)^2,$$
$$e_8(\lambda) = (\lambda - 2)(\lambda - 5)^2,$$
$$e_7(\lambda) = \lambda - 2,$$

mientras que $e_6(\lambda) = \ldots = e_1(\lambda) = 1$.

Ahora que hemos aprendido a escribir inmediatamente la forma canónica de su matriz característica, partiendo de la forma dada de Jordan J, se puede demostrar el siguiente **teorema:**

Dos matrices de Jordan son semejantes si, y sólo si, éstas constan de unas mismas mallas de Jordan, o sea, que solamente pueden diferenciarse en el orden de colocación de estas mallas a lo largo de la diagonal principal.

En efecto, la tabla de polinomios (7) se determinaba completamente por el conjunto de las mallas de Jordan de la matriz de Jordan J, y en ella de ningún modo se reflejaba la colocación de las mallas de Jordan a lo largo de la diagonal principal de esta matriz. De aquí se deduce que, si las matrices de Jordan J y J' poseen una misma colección de mallas de Jordan, a éstas corresponde una misma tabla de polinomios (7), y por esto, unos mismos polinomios (8). Por lo tanto, las matrices características $J - \lambda E$ y $J' - \lambda E$ tienen unos mismos factores invariantes, o sea, son equivalentes, y por lo tanto, las matrices J y J' son semejantes.

Recíprocamente, si las matrices de Jordan son semejantes, sus matrices características tienen iguales factores invariantes. Supongamos que los polinomios (8) para $j = 1, 2, \ldots, q$, son los factores invariantes de éstos, que son distintos de la unidad. Pero, con los polinomios (8) se restablece la tabla de los polinomios (7). Más exactamente, los polinomios (8) se descomponen en productos de potencias de factores lineales, puesto que, como ya se ha demostrado, para cualquier matriz de Jordan los factores invariantes de la matriz característica poseen esta misma propiedad. Precisamente, la tabla (7) consta de las potencias máximas de los factores lineales en que se descomponen los polinomios (8). Finalmente, con la tabla (7) se restablecen las mallas de Jordan de las matrices iniciales de Jordan, pues, a cada polinomio $(\lambda - \lambda_i)^{k_{ij}}$ en la tabla (7) corresponde una malla de Jordan de orden k_{ij}, correspondiente al número λ_i. Con esto, queda demostrado que las matrices J y J' constan de unas mismas mallas de Jordan y que se pueden diferenciar solamente por la colocación de éstas.

En particular, de este teorema se deduce que, *una matriz de Jordan que es semejante a una matriz diagonal, es también diagonal, y que dos matrices diagonales son semejantes si, y sólo si, se diferencian*

entre sí en una permutación de los números que figuran en la diagonal principal.

Reducción de una matriz a la forma normal de Jordan. Si una matriz A con elementos del campo P *se reduce a la forma normal de Jordan*, o sea, es semejante a una matriz de Jordan, entonces, como esto se deduce del teorema demostrado más arriba, *la forma normal de Jordan se determina por la matriz A unívocamente, salvo el orden de colocación de las mallas en la diagonal principal.* La condición para que la matriz A permita tal reducción se indica en el siguiente teorema, cuya demostración nos proporciona a la vez un método práctico para hallar la matriz de Jordan que es semejante a la matriz A, si tal matriz de Jordan existe. Obsérvese en este caso, que **la reducción en el campo P** significa que todos los elementos de la matriz con la que se efectúa la transformación pertenecen al campo P.

Una matriz A con elementos del campo P se reduce en este campo a la forma normal de Jordan cuando, y sólo cuando, todas las raíces características de la matriz A pertenecen al mismo campo fundamental P.

En efecto, si la matriz A es semejante a una matriz de Jordan J, entonces, estas dos matrices tienen unas mismas raíces características. Pero las raíces características de la matriz J se hallan sin dificultad alguna; como el determinante de la matriz $J - \lambda E$ es igual al producto de sus elementos que están en la diagonal principal, el polinomio $|J - \lambda E|$ se descompone sobre el campo P en factores lineales y los números que están en la diagonal principal de la matriz J, y sólo éstos, son sus raíces.

Recíprocamente, supongamos que todas las raíces características de la matriz A pertenecen al mismo campo P. Si

$$e_{n-q+1}(\lambda), \ \ldots, \ e_{n-1}(\lambda), \ e_n(\lambda), \tag{10}$$

son los factores invariantes de la matriz $A - \lambda E$ distintos de 1, entonces,

$$|A - \lambda E| = (-1)^n e_{n-q+1}(\lambda) \ldots e_{n-1}(\lambda) e_n(\lambda).$$

En efecto, los determinantes de la matriz $A - \lambda E$ y de su matriz canónica sólo pueden diferenciarse entre sí en un factor constante que, en realidad, es igual a $(-1)^n$, puesto que así es el coeficiente superior del polinomio característico $|A - \lambda E|$. Por lo tanto, entre los polinomios (10) no hay iguales a cero, la suma de los grados de estos polinomios es igual a n y todos ellos se descomponen sobre el campo P en factores lineales; esto último es debido a que, por la condición, el polinomio $|A - \lambda E|$ tiene tal descomposición.

Sean (8) las descomposiciones de los polinomios (10) en productos de potencias de factores lineales. Llamemos *divisores elementales del polinomio* e_{n-j+1}, $j = 1, 2, \ldots, q$, a las potencias de distintos binomios lineales, diferentes de la unidad, que figuran en su descom-

posición (8), o sea,

$$(\lambda-\lambda_1)^{k_{1j}},\ (\lambda-\lambda_2)^{k_{2j}},\ \ldots,\ (\lambda-\lambda_t)^{k_{tj}}.$$

A los divisores elementales de todos los polinomios (10) los llamaremos *divisores elementales de la matriz A* y los escribiremos en forma de la tabla (7).

Tomemos ahora una matriz de Jordan J de orden n, formada por mallas de Jordan, definidas del modo siguiente: **a cada divisor elemental $(\lambda-\lambda_i)^{k_{ij}}$ de la matriz A ponemos en correspondencia la malla de Jordan de orden k_{ij} que corresponde al número λ_i.** Es evidente, que los polinomios (10), y sólo éstos, son los factores invariantes de la matriz $J-\lambda E$ distintos de la unidad. Por esto, las matrices $A-\lambda E$ y $J-\lambda E$ son equivalentes y, por consiguiente, la matriz A es semejante a la matriz de Jordan J.

Ejemplo. Sea dada la matriz

$$A=\begin{pmatrix} -16 & -17 & 87 & -108 \\ 8 & 9 & -42 & 54 \\ -3 & -3 & 16 & -18 \\ -1 & -1 & 6 & -8 \end{pmatrix}.$$

Reduciendo la matriz $A-\lambda E$ de un modo ordinario a la forma canónica, obtenemos que los factores invariantes de esta matriz, distintos de la unidad, son los polinomios

$$e_4(\lambda)=(\lambda-1)^2(\lambda+2),$$
$$e_3(\lambda)=\lambda-1.$$

Vemos, pues, que la matriz A se reduce a la forma normal de Jordan incluso en el campo de los numeros racionales. Sus divisores elementales son los polinomios $(\lambda-1)^2$, $\lambda-1$ y $\lambda+2$, por lo cual, la matriz

$$J=\begin{pmatrix} 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -2 \end{pmatrix}$$

es la forma normal de Jordan de la matriz A.

Si quisiéramos hallar la matriz no degenerada que transforma la matriz A en la matriz J, tendríamos que valernos de las indicaciones hechas al fin al del párrafo precedente.

Finalmente, basándose en los resultados anteriores, se puede demostrar la siguiente **condición necesaria y suficiente de reducción de una matriz a la forma diagonal**, condición de la que inmediatamente se desprende el criterio suficiente de reducción a la forma diagonal, demostrado en el § 33.

Una matriz A de orden n con elementos del campo P, se reduce a la forma diagonal si, y sólo si, todas las raíces del último factor invariante $e_n(\lambda)$ de su matriz característica pertenecen al campo P, no teniendo que haber múltiples entre ellas.

En efecto, la reducción de una matriz a la forma diagonal es equivalente a la reducción a una forma de Jordan, en la que las mallas de Jordan sean de orden 1. En otras palabras, todos los divisores elementales de la matriz A tienen que ser polinomios de primer grado. Pero, como todos los factores invariantes de la matriz $A - \lambda E$ son divisores del polinomio $e_n(\lambda)$, esta última condición equivale a que todos los divisores elementales del polinomio $e_n(\lambda)$ sean de grado 1, como se quería demostrar.

§ 62. Polinomio mínimo

Sea dada una matriz cuadrada A de orden n con elementos del campo P. Si

$$f(\lambda) = \alpha_0\lambda^k + \alpha_1\lambda^{k-1} + \ldots + \alpha_{k-1}\lambda + \alpha_k$$

es un polinomio del anillo $P[\lambda]$, la matriz

$$f(A) = \alpha_0A^k + \alpha_1A^{k-1} + \ldots + \alpha_{k-1}A + \alpha_kE$$

se llama *valor* del polinomio $f(\lambda)$ para $\lambda = A$; advirtamos que, en este caso, el término independiente del polinomio $f(\lambda)$ se multiplica por la potencia cero de la matriz A, o sea, por la matriz unidad E.

Fácilmente se comprueba que, si

$$f(\lambda) = \varphi(\lambda) + \psi(\lambda),$$

o si

$$f(\lambda) = u(\lambda)\,v(\lambda),$$

entonces,

$$f(A) = \varphi(A) + \psi(A)$$

o, respectivamente,

$$f(A) = u(A)\,v(A).$$

Si la matriz A anula al polinomio $f(\lambda)$, o sea, si

$$f(A) = 0,$$

la matriz A se llamará *raíz matricial*, o bien, cuando esto no dé lugar a confusiones, se llamará simplemente *raíz* del polinomio $f(\lambda)$.

Toda matriz A es raíz de un polinomio no nulo.

En efecto, sabemos que todas las matrices cuadradas de orden n forman sobre el campo P un espacio vectorial de n^2 dimensiones.

De aquí se deduce, que el sistema de $n^2 + 1$ matrices

$$A^{n^2},\ A^{n^2-1},\ \ldots,\ A,\ E,$$

es linealmente dependiente sobre el campo P, o sea, en P existen unos elementos $\alpha_0, \alpha_1, \ldots, \alpha_{n^2}, \alpha_{n^2+1}$, no simultáneamente iguales a cero, tales, que

$$\alpha_0 A^{n^2} + \alpha_1 A^{n^2-1} + \ldots + \alpha_{n^2} A + \alpha_{n^2+1} E = 0.$$

Por lo tanto, resulta que la matriz A es raíz del polinomio no nulo

$$\varphi(\lambda) = \alpha_0 \lambda^{n^2} + \alpha_1 \lambda^{n^2-1} + \ldots + \alpha_{n^2}\lambda + \alpha_{n^2-1},$$

cuyo grado no es superior a n^2.

La matriz A también es raíz de algunos polinomios cuyos coeficientes superiores son iguales a la unidad: es suficiente tomar cualquier polinomio distinto de cero que se anule por la matriz A, y dividirlo por su coeficiente superior. El polinomio de menor grado con el coeficiente superior igual a 1 que se anula por la matriz A, se llama *polinomio mínimo de la matriz A*. Obsérvese, que *el polinomio mínimo de la matriz A se determina unívocamente*, puesto que la diferencia de dos polinomios de éstos sería de menor grado que cada uno de los mismos y se anularía también por la matriz A.

Todo polinomio $f(\lambda)$ que se anula por la matriz A, es divisible por el polinomio mínimo $m(\lambda)$ de esta matriz.

En efecto, si

$$f(\lambda) = m(\lambda)\, q(\lambda) + r(\lambda),$$

donde el grado de $r(\lambda)$ es menor que el grado de $m(\lambda)$, se tiene

$$f(A) = m(A)\, q(A) + r(A)$$

y como $f(A) = m(A) = 0$, resulta, $r(A) = 0$, lo cual contradice a la definición del polinomio mínimo.

Demostremos ahora el siguiente **teorema**:

El polinomio mínimo de una matriz A coincide con el último factor invariante $e_n(\lambda)$ de la matriz característica $A - \lambda E$.

Demostración. Conservando las notaciones y aplicando los resultados del § 59, se puede escribir la igualdad

$$(-1)^n \,|A - \lambda E| = d_{n-1}(\lambda)\, e_n(\lambda). \qquad (1)$$

En particular, de aquí se deduce que los polinomios $e_n(\lambda)$ y $d_{n-1}(\lambda)$ no son nulos. Designemos ahora con $B(\lambda)$ la matriz adjunta a la matriz $A - \lambda E$ (véase el § 14),

$$B(\lambda) = (A - \lambda E)^*.$$

Como se deduce del § 14 (igualdad (3)), se cumple la igualdad

$$(A-\lambda E)\,B(\lambda)=|A-\lambda E|\,E. \qquad (2)$$

Por otra parte, como los menores de $(n-1)$*-ésimo orden de la* matriz $A-\lambda E$, tomados con los signos más o menos, y sólo éstos, son elementos de la matriz $B(\lambda)$, y el polinomio $d_{n-1}(\lambda)$ es el máximo común divisor de todos estos menores, se tiene:

$$B(\lambda)=d_{n-1}(\lambda)\,C(\lambda), \qquad (3)$$

en donde **el máximo común divisor de los elementos de la matriz** $C(\lambda)$ **es igual a 1.**

Pero, de las igualdades (2), (3) y (1), se deduce la igualdad

$$(A-\lambda E)\,d_{n-1}(\lambda)\,C(\lambda)=(-1)^n\,d_{n-1}(\lambda)\,e_n(\lambda)\,E.$$

Esta igualdad se puede simplificar por el factor no nulo $d_{n-1}(\lambda)$, lo cual se deduce de la siguiente observación general: si $\varphi(\lambda)$ es un polinomio no nulo, y $D(\lambda)=(d_{ij}(\lambda))$ es una λ-matriz no nula, donde suponemos que $d_{st}(\lambda)\neq 0$, entonces, en la matriz $\varphi(\lambda)\,D(\lambda)$, en el lugar $(s,\ t)$ figurará el elemento $\varphi(\lambda)\,d_{st}(\lambda)$, distinto de cero. Por lo tanto,

$$(A-\lambda E)\,C(\lambda)=(-1)^n\,e_n(\lambda)\,E,$$

de donde

$$e_n(\lambda)\,E=(\lambda E-A)\,[(-1)^{n+1}\,C(\lambda)] \qquad (4)$$

Esta igualdad muestra que el residuo de la división «a la izquierda» de la λ-matriz que figura en el primer miembro por el binomio $\lambda E-A$, es igual a cero. Sin embargo, del lema demostrado al final del § 60 se deduce que este residuo es igual a la matriz $e_n(A)\,E=e_n(A)$. En efecto, la matriz $e_n(\lambda)\,E$ se puede escribir en forma de un λ-polinomio matricial cuyos coeficientes son matrices escalares, o sea, son conmutables con la matriz A. Por lo tanto,

$$e_n(A)=0,$$

o sea, **el polinomio** $e_n(\lambda)$ **verdaderamente se anula por la matriz** A.

De aquí se deduce, que el polinomio $e_n(\lambda)$ es divisible por el polinomio mínimo $m(\lambda)$ de la matriz A,

$$e_n(\lambda)=m(\lambda)\,q(\lambda). \qquad (5)$$

Está claro, que **el coeficiente superior del polinomio q** (λ) **es igual a la unidad.**

Como $m(A)=0$, de nuevo, en virtud del mismo lema del § 60, el residuo de la división «a la izquierda» de la λ-matriz $m(\lambda)\,E$ por el binomio $\lambda E-A$, es igual a cero, o sea,

$$m(\lambda)\,E=(\lambda E-A)\,Q(\lambda) \qquad (6)$$

Las igualdades (5), (4) y (6) nos llevan a la igualdad

$$(\lambda E - A)[(-1)^{n+1} C(\lambda)] = (\lambda E - A)[Q(\lambda) q(\lambda)].$$

Ambos miembros de esta igualdad se pueden simplificar por el factor común $\lambda E - A$, pues, el coeficiente superior E de este λ-polinomio matricial es una matriz no degenerada. Por lo tanto,

$$C(\lambda) = (-1)^{n+1} Q(\lambda) q(\lambda).$$

Recordemos, sin embargo, que el máximo común divisor de los elementos de la matriz $C(\lambda)$ es igual a 1. Por esto, el polinomio $q(\lambda)$ tiene que ser de grado cero, y como su coeficiente superior es igual a 1, resulta, $q(\lambda) = 1$. Por lo tanto, en virtud de (5),

$$e_n(\lambda) = m(\lambda),$$

que es lo que se quería demostrar.

Como, en virtud de (1), el polinomio característico de la matriz A es divisible por el polinomio $e_n(\lambda)$, del teorema que acabamos de demostrar se desprende el siguiente

Teorema de Hamilton-Cayley. *Toda matriz es raíz de su polinomio característico.*

Polinomio mínimo de una transformación lineal. Demostremos primero la siguiente proposición:

Si las matrices A y B son semejantes y la matriz A anula al polinomio $f(\lambda)$, entonces, la matriz B también anula al mismo.

En efecto, sea

$$B = C^{-1}AC.$$

Si

$$f(\lambda) = \alpha_0\lambda^k + \alpha_1\lambda^{k-1} + \ldots + \alpha_{k-1}\lambda + \alpha_k,$$

se tiene

$$\alpha_0 A^k + \alpha_1 A^{k-1} + \ldots + \alpha_{k-1}A + \alpha_k E = 0.$$

Transformando ambos miembros de esta igualdad con la matriz C, obtenemos:

$$C^{-1}(\alpha_0 A^k + \alpha_1 A^{k-1} + \ldots + \alpha_{k-1}A + \alpha_k E)\,C =$$
$$= \alpha_0 (C^{-1}AC)^k + \alpha_1 (C^{-1}AC)^{k-1} + \ldots + \alpha_{k-1}(C^{-1}AC) + \alpha_k E =$$
$$= \alpha_0 B^k + \alpha_1 B^{k-1} + \ldots + \alpha_{k-1}B + \alpha_k E = 0,$$

o sea, $f(B) = 0$.

De aquí se deduce, que *las matrices semejantes poseen un mismo polinomio mínimo.*

Supongamos ahora que φ es una transformación lineal del espacio lineal de n dimensiones sobre el campo P. Las matrices que determinan esta transformación en distintas bases del espacio, son seme-

jantes entre sí. El polinomio mínimo común de estas matrices se llama *polinomio mínimo de la transformación lineal* φ.

Aplicando las operaciones sobre las transformaciones lineales, introducidas en el § 32, se puede introducir el concepto de *valor* de un polinomio

$$f(\lambda)=\alpha_0\lambda^k+\alpha_1\lambda^{k-1}+\ldots+\alpha_{k-1}\lambda+\alpha_k$$

del anillo $P[\lambda]$ para λ, igual a una transformación lineal φ: este valor será la transformación lineal

$$f(\varphi)=\alpha_0\varphi^k+\alpha_1\varphi^{k-1}+\ldots+\alpha_{k-1}\varphi+\alpha_k\varepsilon,$$

donde ε es la transformación idéntica.

Diremos luego que la transformación lineal φ *anula* al polinomio $F(\lambda)$, si

$$f(\varphi)=\omega,$$

donde ω es la transformación nula.

Teniendo en cuenta la relación existente entre las operaciones sobre las transformaciones lineales y sobre las matrices, el lector demostrará sin dificultad alguna, que *el polinomio mínimo de una transformación lineal* φ *es el polinomio de menor grado con el coeficiente superior* 1, *determinado unívocamente, que se anula por la transformación* φ. Después de esto, los resultados obtenidos anteriormente y, en particular, el teorema de Hamilton-Cayley, se pueden enunciar de nuevo en términos de transformaciones lineales.

CAPITULO XIV

GRUPOS

§ 63. Definición y ejemplos de grupos

Los anillos y los cuerpos, que desempeñaron un papel tan grande en los capítulos anteriores, son sistemas algebraicos de dos operaciones independientes: adición y multiplicación. Sin embargo, en diversas ramas de las matemáticas y en sus aplicaciones, frecuentemente se encuentran tales sistemas algebraicos, en los que está definida una sola operación algebraica. Así, pues, limitándonos por ahora a los ejemplos que ya aparecieron en nuestro libro, señalemos, que en el conjunto de las sustituciones de grado n (véase el § 3), solamente habíamos definido una operación: la multiplicación. Por otra parte, en la definición del espacio vectorial (§ 8) está incluida la suma de vectores, mientras que el producto de vectores no había sido definido (señalemos, que el producto de un vector por un número no satisface a la definición de operación algebraica dada en el § 44).

Un tipo importante de sistemas algebraicos con una operación son los grupos. Este concepto posee un campo extraordinariamente amplio de aplicaciones y representa el objeto de una gran ciencia independiente, de la teoría de los grupos. El capítulo presente puede considerarse como introducción a la teoría de los grupos: en él se expondrán las nociones elementales sobre los grupos, cuyo conocimiento es necesario para cada matemático; el capítulo se terminará con la exposición de un teorema menos elemental.

De acuerdo a la teoría general de los grupos, convengamos en llamar *multiplicación* a la operación algebraica considerada y en emplear los símbolos correspondientes. Recordemos (véase el § 44), que **se supone que siempre es posible la operación algebraica, y que ésta es univalente:** para cualquier par de elementos a y b del conjunto considerado, existe el producto ab y representa un elemento unívocamente determinado de este conjunto.

Se llama *grupo* a un conjunto G con una operación algebraica, que es asociativa (aunque no necesariamente conmutativa), y para la que existe además la operación inversa.

Como la operación en el grupo puede ser no conmutativa, la existencia de la operación inversa significa lo siguiente: para cualquier par de elementos a y b de G, existe en G un elemento x y un elemento y, **unívocamente determinados**, tales que

$$ax = b, \quad ya = b.$$

Si el grupo G se compone de un número finito de elementos, se denomina *grupo finito*, y el número de sus elementos, se llama *orden* del grupo. Si la operación definida en el grupo es conmutativa, G se denomina grupo *conmutativo* o *abeliano*.

Señalemos las consecuencias elementales de la definición de grupo. Basándose en los razonamientos expuestos ya en el § 44, se puede afirmar que la ley asociativa nos permite hablar de un modo unívoco del *producto de un número finito cualquiera de elementos del grupo*, dados en un orden determinado (ya que la operación en el grupo puede ser no conmutativa).

Veamos las consecuencias de la existencia de la operación inversa.

Supongamos que en el grupo G se ha dado un elemento arbitrario a. De la definición del grupo se deduce la existencia en G de un elemento e_a, unívocamente determinado, tal que $ae_a = a$; por consiguiente, este elemento desempeña el papel de la unidad al multiplicar el elemento a por él a la derecha. Si b es otro elemento cualquiera del grupo G, y si y es el elemento del grupo que satisface a la igualdad $ya = b$, cuya existencia se deduce de la definición del grupo, se tiene:

$$b = ya = y(ae_a) = (ya)e_a = be_a.$$

Por lo tanto, el elemento e_a desempeña el papel de unidad a la derecha con respecto a todos los elementos del grupo G y no sólo con respecto al elemento inicial a; por eso, lo designaremos mediante e'. De la unicidad, que forma parte de la definición de la operación inversa, se deduce la unicidad de este elemento.

De este mismo modo se puede demostrar la existencia en G y la unicidad de un elemento e'' que satisfaga a la condición $e''a = a$ para todos los elementos a de G. En realidad, los elementos e' y e'' coinciden, puesto que de las igualdades $e''e' = e''$ y $e''e' = e'$ se deduce que $e'' = e'$. De esta manera, queda demostrado que *en cada grupo G existe un elemento e, unívocamente determinado, que satisface a la condición:*

$$ae = ea = a$$

para todos los elementos a de G. Este elemento se llama *unidad* del grupo G y se designa ordinariamente con el símbolo 1.

Para cada elemento dado a, de la definición del grupo se deduce, la existencia y unicidad de unos elementos a' y a'' tales, que

$$aa' = 1, \quad a''a = 1.$$

En la realidad, los elementos a' y a'' coinciden: de las igualdades

$$a''aa' = a''(aa') = a''\cdot 1 = a'',$$
$$a''aa' = (a''a)a' = 1\cdot a' = a',$$

se deduce que $a'' = a'$. Este elemento se llama *inverso* del elemento a y se designa con la notación a^{-1}, de modo que

$$aa^{-1} = a^{-1}a = 1.$$

Por lo tanto, *cada elemento del grupo posee un elemento inverso, unívocamente determinado.*

De las últimas igualdades se deduce, que el mismo elemento a sirve de inverso para el elemento a^{-1}. Es fácil observar también, que el inverso del producto de unos cuantos elementos es el producto de los elementos inversos de los factores y, además, tomados en orden inverso:

$$(a_1a_2 \ldots a_{n-1}a_n)^{-1} = a_n^{-1}a_{n-1}^{-1} \ldots a_2^{-1}a_1^{-1}.$$

Por fin, el elemento inverso de la unidad es la unidad misma.

La prueba para averiguar si un conjunto dado con una operación es grupo o no, se facilita sumamente por el hecho de que en la definición de grupo la demanda del cumplimiento de la operación inversa se puede sustituir por la suposición de la existencia de la unidad y de los elementos inversos y, además, sólo por un lado (por ejemplo, por la derecha) y sin suponer la unicidad de ellos. Esto se deduce del siguiente **teorema:**

Un conjunto G con una operación asociativa es grupo, si en él existe por lo menos un elemento e que posee la propiedad:

$$ae = a \text{ para todos los elementos } a \text{ de } G,$$

y si entre todos los elementos unidades a la derecha existe por lo menos un elemento e_0 tal, que con respecto a él cada elemento a de G posee por lo menos un elemento inverso a la derecha a^{-1}:

$$aa^{-1} = e_0.$$

Demostración. Sea a^{-1} uno de los elementos inversos a la derecha de a. Entonces,

$$aa^{-1} = e_0 = e_0e_0 = e_0aa^{-1},$$

o sea, $aa^{-1} = e_0aa^{-1}$. Multiplicando a la derecha ambos miembros de esta igualdad por uno de los elementos que son inversos a la derecha de a^{-1}, obtenemos, $ae_0 = e_0\, ae_0$, de donde $a = e_0a$, puesto que e_0 es una unidad a la derecha de G. Por lo tanto, resulta que el elemento e_0 es también una unidad a la izquierda de G. Si ahora e_1 es una unidad a la derecha arbitraria y e_2 es una unidad a la izquierda

arbitraria, de las igualdades

$$e_2 e_1 = e_1 \text{ y } e_2 e_1 = e_2$$

se deduce que $e_1 = e_2$, o sea, que cualquier unidad a la derecha es igual a cualquier unidad a la izquierda. Queda, pues, demostrada la existencia y unicidad en el conjunto G del elemento unidad, que lo indicaremos, como anteriormente, mediante 1.

Luego,

$$a^{-1} = a^{-1} \cdot 1 = a^{-1} a a^{-1},$$

es decir, $a^{-1} = a^{-1}\, a a^{-1}$, donde a^{-1} es uno de los elementos inversos a la derecha de a. Multiplicando a la derecha ambos miembros de *la última igualdad por uno de los elementos inversos a la derecha* de a^{-1}, obtenemos, $1 = a^{-1}\, a$, o sea, que el elemento a^{-1} sirve también de elemento inverso a la izquierda de a. Si ahora a_1^{-1} es un elemento inverso a la derecha arbitrario de a, y a_2^{-1} es un elemento inverso a la izquierda arbitrario del mismo, de las igualdades

$$a_2^{-1} a a_1^{-1} = (a_2^{-1} a)\, a_1^{-1} = a_1^{-1},$$
$$a_2^{-1} a a_1^{-1} = a_2^{-1}\, (a a_1^{-1}) = a_2^{-1}$$

e deduce que $a_1^{-1} = a_2^{-1}$, es decir, se deduce la existencia y la unicidad, para cada elemento a de G, del elemento inverso a^{-1}.

Ahora es fácil mostrar que el conjunto G es grupo. En efecto, como bien se obverva, las ecuaciones $ax = b$, $ya = b$ se satisfacen con los elementos

$$x = a^{-1} b, \quad y = b a^{-1}.$$

La unicidad de estas soluciones se deduce de que si, por ejemplo, $ax_1 = ax_2$, multiplicando a la izquierda ambos miembros de esta igualdad por a^{-1}, obtenemos $x_1 = x_2$. El teorema queda demostrado.

Ya nos hemos encontrado unas cuantas veces con el concepto de isomorfismo: para los anillos, para los espacios lineales, para los espacios euclídeos. Este concepto puede ser definido también para los grupos y desempeña en la teoría de los mismos un papel tan importante como en la teoría de los anillos. Se dice que los grupos G y G' son *isomorfos*, si se puede establecer entre ellos una correspondencia biunívoca tal, que para cualquier par de elementos a y b de G y para sus correspondientes elementos a' y b' de G', al producto ab corresponde el producto $a'b'$. Del mismo modo que en el § 46 (para el cero y para el elemento opuesto del anillo), se puede demostrar que, en la correspondencia isomorfa de los grupos G y G', a la unidad del grupo G corresponde la unidad del grupo G', y si al elemento a de G le corresponde el elemento a' de G', al elemento a^{-1} le corresponderá el elemento a'^{-1}.

Pasando a examinar **ejempolos de grupos,** señalemos que, si la operación en el grupo se llamase *suma*, la unidad del grupo se llamaría *cero* y se indicaría con la notación 0, y en lugar de elemento inverso diríamos *elemento opuesto* y lo indicaríamos mediante $-a$.

Como primer ejemplo de grupo, anotemos que, *respecto a la suma, cualquier anillo* (*y, en particular, un cuerpo*) *representa un grupo, y además, abeliano*; éste es el llamado *grupo aditivo del anillo*. Esta observación proporciona inmediatamente una gran cantidad de ejemplos concretos de grupos, y entre ellos: el grupo aditivo de números enteros, el grupo aditivo de números pares, los grupos aditivos de números racionales, de números reales, de números complejos, etc, etc. Señalemos, que *los grupos aditivos de números enteros y de números pares son isomorfos entre sí*, a pesar de que el segundo forma sólo una parte del primero: la transformación que pone en correspondencia a cada número entero k el número par $2k$, es biunívoca y, como fácilmente se puede comprobar, representa una transformación isomorfa del primero de los grupos nombrados sobre el segundo.

Ningún anillo es grupo respecto a la multiplicación, puesto que no siempre se cumple la operación inversa, que es la división. No cambia el asunto al pasar de un anillo arbitrario a un cuerpo, puesto que en éste se mantiene sin cumplir la división por cero. Examinemos, sin embargo, el conjunto de todos los elementos del cuerpo diferentes de cero. Como el campo no contiene divisores de cero, es decir, que el producto de dos elementos diferentes de cero también es diferente de cero, la multiplicación representa una operación algebraica para el conjunto considerado, que es asociativa y conmutativa, siendo posible ya la división sin salir fuera de los límites de este conjunto. Por lo tanto, *el conjunto de todos los elementos diferentes de cero de cualquier campo representa un grupo abeliano*; éste se llama *grupo multiplicativo* del campo. Ejemplos concernientes a esto son: los grupos multiplicativos de números racionales, de números reales, de números complejos.

Es evidente que, respecto a la multiplicación, todos los números reales positivos forman grupo. *Este grupo es isomorfo al grupo aditivo de todos los números reales*: poniendo en correspondencia a cada número positivo a el número real $\ln a$, obtenemos una aplicación biyectiva del primero de los grupos sobre el segundo, que representa un isomorfismo en vista de la igualdad,

$$\ln (ab) = \ln a + \ln b.$$

Tomemos, ahora, en el campo de los números complejos, el conjunto de las raíces n-ésimas de la unidad. En el § 19 se había demostrado que el producto de dos raíces n-ésimas de la unidad, así como el número recíproco de la raíz n-ésima de la unidad, pertenecen al mismo conjunto considerado de números. Como la unidad también

pertenece, naturalmente, a este conjunto, y como la multiplicación de cualesquiera números complejos es asociativa y conmutativa, obtenemos que *las raíces n-ésimas de la unidad forman un grupo abeliano respecto a la multiplicación; este grupo es finito y de orden n. Por lo tanto, para cualquier número natural n, existen grupos finitos de orden n.*

El grupo (respecto a la multiplicación) de las raíces n-ésimas de la unidad es isomorfo al grupo aditivo del anillo Z_n construido en el § 45. En efecto, si ε es una raíz primitiva de orden n de la unidad, todos los elementos del primero de los grupos considerados tienen la forma ε^k, $k = 0, 1, \ldots, n-1$. Si ponemos en correspondencia a cada número ε^k el elemento C_k del anillo Z_n, o sea, la clase de números enteros cuyos residuos, al dividirlos por n, son iguales a k, obtenemos una correspondencia de isomorfismo entre los grupos considerados: si $0 \leqslant k \leqslant n-1$, $0 \leqslant l \leqslant n-1$ y si $k + l = nq + r$, donde $0 \leqslant r \leqslant n-1$, y q es igual a 0 ó a 1, entonces, $\varepsilon^k \cdot \varepsilon^l = \varepsilon^r$ y, a la vez, $C_k + C_l = C_r$.

Es oportuno señalar ahora unos cuantos ejemplos de conjuntos numéricos que no forman grupo. Así, el conjunto de todos los números enteros no forma grupo respecto a la multiplicación, el conjunto de todos los números reales positivos no forma grupo respecto a la suma, el conjunto de todos los números impares no forma grupo respecto a la suma, el conjunto de todos los números reales negativos no forma grupo respecto a la multiplicación. No representa dificultad alguna la comprobación de todas estas afirmaciones.

Naturalmente, todos los grupos numéricos examinados anteriormente son abelianos. Los espacios lineales sirven de ejemplos de grupos abelianos que no están formados por números: como se deduce de su definición (véase los §§ 29, 47), *todo espacio lineal sobre un cuerpo arbitrario P es grupo abeliano respecto a la operación de la suma.*

Veamos algunos ***ejemplos de grupos no conmutativos.***

El conjunto de todas las matrices de orden n sobre un campo P no representa grupo respecto a la operación de multiplicar, ya que no se cumple la condición de existencia del elemento inverso. Sin embargo, *si nos limitamos sólo a las matrices que no son degeneradas, se obtiene ya un grupo.* En efecto, como sabemos, el producto de dos matrices no degeneradas es una matriz no degenerada, la matriz unidad tampoco es degenerada, toda matriz no degenerada posee matriz inversa, que tampoco es degenerada, y, por fin, la ley asociativa, cumpliéndose para todas las matrices, se cumple, particularmente, para las matrices no degeneradas. Por consiguiente, se puede hablar del *grupo de las matrices no degeneradas de orden n sobre el cuerpo P*, tomando por operación en el grupo el producto de las matrices; este grupo no es conmutativo para $n \geqslant 2$.

El producto de sustituciones, definido en el § 3, da lugar a ejemplos muy importantes de grupos finitos no conmutativos. Ya sabemos que, en el conjunto de todas las sustituciones de grado n, la multiplicación representa una operación algebraica, que es, además, asociativa, aunque para $n \geqslant 3$ no es conmutativa; también sabemos que la sustitución idéntica E sirve de unidad en esta multiplicación y que para cualquier sustitución existe la sustitución inversa. Por lo tanto, *el conjunto de las sustituciones de grado n forma grupo respecto a la multiplicación, que es además finito y de orden* $n!$. Este se llama *grupo simétrico de grado* n, y para $n \geqslant 3$ no es conmutativo.

En lugar de examinar el conjunto de sustituciones de grado n, consideremos ahora solamente el conjunto de las sustituciones pares, compuesto, como ya sabemos, de $\frac{1}{2}n!$ elementos. Aplicando el teorema demostrado en el § 3, según el cual la paridad de la sustitución coincide con la paridad del número de trasposiciones que forman parte en cualquiera de las descomposiciones de esta sustitución en producto de trasposiciones, se obtiene, que *el producto de dos sustituciones pares es una sustitución par*; en efecto, la descomposición de AB en forma de un producto de trasposiciones se obtiene yuxtaponiendo las descomposiciones correspondientes de A y B. Ya se sabe que es asociativa la multiplicación de sustituciones; es evidente, que la sustitución idéntica es par. Por fin, es par la sustitución A^{-1}, si es par la sustitución A; esto es debido aunque sólo sea al hecho de que las expresiones de estas sustituciones se pueden obtener una de otra permutando de sitio las filas superior e inferior, o sea, que ellas contienen igual número de inversiones. Por consiguiente, *el conjunto de las sustituciones pares de n grado representa un grupo finito respecto a la multiplicación, de orden* $\frac{1}{2}n!$. Este se llama *grupo alternado de grado* n; es fácil comprobar que este grupo no es conmutativo para $n \geqslant 4$, a pesar de que es conmutativo para $n = 3$.

Los grupos simétrico y alternado desempeñan un gran papel en la teoría de los grupos finitos y también en la teoría de Galois. Señalemos que, por analogía con los grupos alternados, sería imposible construir con las sustituciones impares un grupo respecto a la multiplicación, puesto que el producto de dos sustituciones impares siempre es una sustitución par.

Las diversas ramas de la geometría proporcionan numerosos ejemplos de grupos distintos. Indiquemos un ejemplo sencillo de este género: el conjunto de todas las rotaciones de una esfera alrededor de su centro representa un grupo, pero no conmutativo, si es que llamamos producto de dos rotaciones al resultado de su realización consecutiva.

§ 64. Subgrupos

Un subconjunto A de un grupo G se llama *subgrupo* de éste si él mismo representa un grupo respecto a la operación definida en el grupo G.

Para verificar que el subconjunto A del grupo G forma un subgrupo de este grupo, es suficiente comprobar: 1) si contiene A el producto de dos elementos cualesquiera de A; 2) si contiene A, junto con cada uno de sus elementos, el elemento inverso. En efecto, del cumplimiento de la ley asociativa en el grupo G se deduce su cumplimiento para los elementos de A, y la pertenencia de la unidad del grupo G a A es consecuencia de 2) y 1).

Muchos de los grupos señalados en el párrafo anterior representan subgrupos de otros grupos indicados allí mismo. Así, el grupo aditivo de los números pares representa un subgrupo del grupo aditivo de los números enteros, y este último a su vez es un subgrupo del grupo aditivo de los números racionales. Todos estos grupos, como en general los grupos aditivos de números representan subgrupos del grupo aditivo de los números complejos. El grupo multiplicativo de los números reales positivos representa un subgrupo del grupo multiplicativo de todos los números reales diferentes de cero. El grupo alternado de grado n es un subgrupo del grupo simétrico del mismo grado.

Subrayemos, que la condición que figura en la definición de subgrupo, de que **el subconjunto A del grupo G sea grupo respecto a la operación definida en el grupo G**, es esencial. Así, el grupo multiplicativo de los números reales positivos **no representa** un subgrupo del grupo aditivo de todos los números reales, a pesar de que el primer conjunto está contenido en el segundo como subconjunto.

Si en el grupo G se han tomado los subgrupos A y B, su intersección $A \cap B$, es decir, el conjunto de los elementos pertenecientes a A y a B, también es un subgrupo del grupo G.

En efecto, si los elementos x e y pertenecen a la intersección $A \cap B$, estos pertenecen al subgrupo A, y por eso, el producto xy y el elemento inverso x^{-1} también pertenecen a A. Por las mismas razones, los elementos xy y x^{-1} pertenecen también al subgrupo B, y por eso, éstos pertenecen también a $A \cap B$.

Como fácilmente se ve, el resultado obtenido no sólo es justo para dos grupos, sino que también lo es para un número cualquiera de subgrupos, finito e incluso infinito.

El subconjunto del grupo G formado por el solo elemento 1, representa, evidentemente, un subgrupo de este grupo; este subgrupo, que está contenido en cualquier otro subgrupo del grupo G, se llama *subgrupo unidad* del grupo G. Por otra parte, el mismo grupo G representa uno de sus subgrupos.

Los llamados **subgrupos cíclicos** sirven de ejemplos interesantes de subgrupos. Introduzcamos primero el concepto de **potencia** de un elemento a de un grupo G. Siendo n un número natural arbitrario, el producto de n elementos iguales al elemento a se llama *potencia* del elemento a de grado n y se indica mediante a^n. Las *potencias negativas* del elemento a se pueden determinar, bien como elementos del grupo G, inversos a las potencias positivas de este elemento, o bien como el producto de unos cuantos factores, iguales al elemento a^{-1}. En la realidad, estas definiciones coinciden:

$$(a^n)^{-1} = (a^{-1})^n, \quad n > 0. \tag{1}$$

Para la demostración, es suficiente tomar el producto de $2n$ factores, de los cuales, los n primeros sean iguales a a y los demás, a a^{-1}, y efectuar todas las simplificaciones. El elemento igual a ambos miembros de la igualdad (1), se indicará mediante a^{-n}. Convengamos, por fin, en entender por la *potencia cero* a^0 del elemento a, el elemenmento **1**.

Obsérvese, que si la operación en el grupo G se llama suma, en lugar de las potencias del elemento a se debe hablar de los *múltiplos* de este elemento, escribiéndolos mediante ka.

Fácilmente se comprueba que en cualquier grupo G, para las potencias de cualquier elemento a con cualesquiera exponentes m y n, positivos, negativos o ceros, se verifican las igualdades:

$$a^n \cdot a^m = a^m \cdot a^n = a^{n+m}, \tag{2}$$

$$(a^n)^m = a^{nm}. \tag{3}$$

Designemos con $\{a\}$ el subconjunto del grupo G formado por todas las potencias del elemento a; el mismo elemento a también está incluido en él, representando la primera potencia. *El subconjunto $\{a\}$ es un subgrupo del grupo G*: el producto de elementos de $\{a\}$ pertenece a $\{a\}$, en virtud de (2); el elemento 1, igual a a^0, pertenece a $\{a\}$ y, por fin, $\{a\}$ junto con cada elemento suyo contiene al elemento inverso, puesto que de (3) se deduce la igualdad

$$(a^n)^{-1} = a^{-n}.$$

El subgrupo $\{a\}$ se llama *subgrupo cíclico del grupo G, engendrado por el elemento a*. Como muestra la igualdad (2), este subgrupo siempre es conmutativo, incluso cuando el mismo grupo G no sea conmutativo.

Señalemos, que anteriormente no se había afirmado nunca que todas las potencias del elemento a son diferentes elementos del grupo. Si esto es verdaderamente así, entonces a se llama *elemento de orden infinito*. Sin embargo, supongamos que entre las potencias del elemento a haya algunas iguales, por ejemplo, $a^k = a^l$ siendo $k \neq l$; esto siempre tiene lugar en el caso de grupos finitos, pero puede

ocurrir también en un grupo infinito. Si $k > l$, se tiene

$$a^{k-l} = 1,$$

es decir, existen potencias **positivas** del elemento a que son iguales a la unidad. Supongamos que n es la potencia positiva menor del elemento a, que es igual a la unidad, o sea, que

$$1)\ a^n = 1, \quad n > 0,$$

$$2)\ \text{si } a^k = 1, \quad k > 0, \text{ entonces } k \geqslant n.$$

En este caso, se dice que a es un *elemento de orden finito*, precisamente, *de orden* n.

Es fácil observar que, si el elemento a es de orden finito n, todos los elementos

$$1,\ a,\ a^2,\ \ldots,\ a^{n-1} \tag{4}$$

son diferentes. *Cualquiera otra potencia del elemento a, positiva o negativa, es igual a uno de los elementos* (4). En efecto, si k es un número entero arbitrario, dividiéndolo por n se obtiene:

$$k = nq + r, \quad 0 \leqslant r \leqslant n,$$

y, por eso, en virtud de (2) y (3),

$$a^k = (a^n)^q \cdot a^r = a^r. \tag{5}$$

De esto se deduce que, *si el elemento a es de orden finito n y $a^k = 1$, entonces k se divide por n.* Por otra parte, como

$$-1 = n(-1) + (n-1),$$

para el elemento a de orden finito n,

$$a^{-1} = a^{n-1}.$$

Como el sistema (4) contiene n elementos, de los resultados obtenidos anteriormente se deduce que, *para un elemento a que tiene orden finito, su orden n coincide con el orden (o sea, con el número de los elementos) del subgrupo cíclico $\{a\}$.*

Señalemos, por fin, que todo grupo posee un elemento único de primer orden: éste es el elemento 1. Es evidente, que el subgrupo cíclico $\{1\}$ coincide con el subgrupo unidad.

Grupos cíclicos. Un grupo G se llama *cíclico* si se compone de las potencias de uno de sus elementos a, es decir, que coincide con uno de sus subgrupos cíclicos $\{a\}$; en este caso, a se llama *elemento generador* del grupo G. Es evidente, que todo grupo cíclico es abeliano.

El grupo aditivo de los números enteros sirve de ejemplo de grupo cíclico infinito, pues todo número entero es múltiplo de **1**, es decir, que este número es el elemento generador del grupo considerado; se podría tomar también como elemento generador el número -1.

El grupo multiplicativo de las raíces de grado n de la unidad sirve de ejemplo de grupo finito cíclico de orden n, pues, como se había mostrado en el § 19, todas estas raíces son potencias de una de ellas, que es, precisamente, la raíz primitiva.

El teorema que sigue muestra que, con estos ejemplos se agotan en la realidad todos los grupos cíclicos:

Todos los grupos cíclicos infinitos son isomorfos entre sí; son isomorfos entre sí también todos los grupos cíclicos finitos de un orden dado n.

En efecto, resulta una aplicación biyectiva del grupo cíclico infinito, con el elemento generador a, sobre el grupo aditivo de los números enteros, al hacer corresponder a cada elemento a^k del primer grupo el número k; esta aplicación representa un isomorfismo, puesto que de acuerdo a (2), al multiplicar las potencias del elemento a se suman los exponentes. Si se da un grupo cíclico finito G de orden n, con el elemento generador a, entonces designando con ε una raíz primitiva de grado n de la unidad asociamos a cada elemento a^k del grupo G el número ε^k, $0 \leqslant k < n$. Esto representa una aplicación biyectiva del grupo G sobre el grupo multiplicativo de las raíces de grado n de la unidad, cuyo isomorfismo se deduce de (2) y (5).

Este teorema da la posibilidad de hablar simplemente del *grupo cíclico infinito*, o bien del *grupo cíclico de orden* n.

Demostremos ahora el **teorema** siguiente:

Todo subgrupo de un grupo cíclico es cíclico.

En efecto, sea $G = \{a\}$ un grupo cíclico con el elemento generador a, finito o infinito, y sea A un subgrupo del grupo G. Se puede suponer que A es diferente del subgrupo unidad, pues, en caso contrario, no habría que demostrar nada. Supongamos que a^k es la potencia positiva mínima del elemento a, contenida en A; tal potencia existe, puesto que si A contiene el elemento a^{-s}, $s > 0$, diferente de **1**, contiene también el elemento a^s, inverso de él. Supongamos que A contiene también al elemento a^l, $l \neq 0$, y que l no es divisible por k. Entonces, si $d, d > 0$, es el máximo común divisor de los números k y l, existen unos números enteros u y v tales, que

$$ku + lv = d,$$

y por eso, el subgrupo A tiene que contener al elemento

$$(a^k)^u \cdot (a^l)^v = a^d,$$

pero como por la hipótesis $d < k$, llegamos a una contradicción con la elección del elemento a^k. Con esto, queda demostrado que $A = \{a^k\}$.

Descomposición de un grupo en clases con relación a un subgrupo. Tomando en el grupo G los subconjuntos M y N, por *producto* MN *de ellos* se entiende el conjunto de los elementos del grupo G que se

pueden representar, aunque sólo sea de un modo, en forma de un producto de un elemento de M por un elemento de N. Del cumplimiento de la ley asociativa para la operación en el grupo se deduce *su cumplimiento para la multiplicación de los subconjuntos del grupo*:

$$(MN)P = M(NP).$$

Naturalmente, uno de los conjuntos M, N puede estar compuesto de un solo elemento a. En este caso, se obtiene *el producto aN del elemento por el conjunto o el producto Ma del conjunto por el elemento.*

Supongamos que en el grupo G se ha dado un subgrupo arbitrario A. Si x es un elemento cualquiera de G, el producto xA se llama *clase adjunta a la izquierda del subgrupo A en el grupo G, engendrada por el elemento x**. Es comprensible, que *el elemento x está contenido en la clase adjunta xA*, puesto que el subgrupo A contiene la unidad, y $x \cdot 1 = x$.

Toda clase adjunta a la izquierda es engendrada por cualquiera de sus elementos, es decir, que si el elemento y pertenece a la clase adjunta xA, entonces,

$$yA = xA. \tag{6}$$

En efecto, y se puede representar en la forma

$$y = xa,$$

donde a es un elemento del subgrupo A. Por eso, para cualesquiera elementos a' y a'' de A, se tiene

$$ya' = x(aa'),$$
$$xa'' = y(a^{-1}a''),$$

con lo que queda demostrada la igualdad (6).

De esto se deduce que *dos clases adjuntas a la izquierda cualesquiera del subgrupo A en el grupo G, o coinciden, o no tienen ningún elemento común*. En efecto, si las clases adjuntas xA e yA contienen un elemento común z, se tiene:

$$xA = zA = yA.$$

Por lo tanto, todo el grupo G se descompone en clases adjuntas a la izquierda, disjuntas respecto al subgrupo A. Esta descomposición se llama *descomposición del grupo G en clases a la izquierda respecto del subgrupo A*.

Adviértase que una de las clases adjuntas a la izquierda de esta descomposición coincide con el mismo subgrupo A; esta clase está

* A veces, se llama clase de restos, clase residual o simplemente clase y también cogrupo. Para evitar confusiones, advirtamos, que un cogrupo nunca es un subgrupo, a excepción del cogrupo engendrado por el elemento unidad (o por cualquier elemento del subgrupo A) que coincide con el mismo subgrupo A. (*Nota del T.*).

engendrada por el elemento 1, o, en general, por cualquier elemento a de A, puesto que

$$aA = A.$$

Es obvio, que llamando al producto Ax *clase adjunta a la derecha del subgrupo A en el grupo G, engendrada por el elemento x*, de modo análogo obtendríamos la *descomposición a la derecha del grupo G respecto del subgrupo A*. Naturalmente, para un grupo abeliano, ambas descomposiciones, a la izquierda o a la derecha, respecto de cualquier subgrupo coinciden, es decir, se puede hablar simplemente de la *descomposición del grupo respecto del subgrupo*.

Así, pues, la descomposición del grupo aditivo de los números enteros con respecto del subgrupo de los números que son múltiplos del número k, se compone de k clases residuales distintas, engendradas por los números $0, 1, 2, \ldots, k-1$, respectivamente. En este caso, en la clase residual, engendrada por el número l, $0 \leqslant l \leqslant k-1$, están comprendidos todos los números que al ser divididos por k dan el resto l.

Cuando el grupo no es conmutativo, sus descomposiciones respecto de un subgrupo pueden ser distintas.

Veamos, por ejemplo, el grupo simétrico de 3er grado S_3, donde, de acuerdo al § 3, se escribirán sus elementos mediante ciclos. Tomemos en calidad de subgrupo A el subgrupo cíclico engendrado por el elemento (12); este subgrupo consta de la sustitución idéntica y de la sustitución (12) misma. Las otras clases adjuntas a la izquierda son: la clase $(13) \cdot A$, que se compone de las sustituciones (13) y (132) y la clase $(23) \cdot A$, que se compone de las sustituciones (23) y (123). Por otra parte, las clases adjuntas a la derecha relativas al subgrupo A son: el mismo subgrupo A, la clase $A \cdot (13)$, compuesta de las sustituciones (13) y (123), y la clase $A \cdot (23)$, compuesta de las sustituciones (23) y (132). Vemos, pues, que en este caso, la descomposición en clases a la derecha se diferencia de la descomposición en clases a la izquierda.

En el caso de grupos finitos, la existencia de descomposiciones de un grupo en clases respecto de un subgrupo nos lleva al siguiente teorema importante:

Teorema de Lagrange. *En todo grupo finito, el orden de cualquier subgrupo es un divisor del orden del mismo grupo.*

En efecto, supongamos que en el grupo finito G de orden n se haya dado un subgrupo A de orden k. Consideremos la descomposición del grupo G en clases a la izquierda respecto del subgrupo A. Supongamos que ésta consta de j clases; el número j se llama *índice* del subgrupo A en el grupo G. Cada clase adjunta a la izquierda xA consta de k elementos, exactamente, puesto que si

$$xa_1 = xa_2,$$

donde a_1 y a_2 son elementos de A, entonces, $a_1 = a_2$. Por lo tanto,

$$n = kj, \tag{7}$$

que es lo que se quería demostrar.

Como el orden de un elemento coincide con el orden de su subgrupo cíclico, del teorema de Lagrange se deduce que *el orden de cada elemento de un grupo finito es divisor del orden del grupo.*

Del teorema de Lagrange se deduce también, que *todo grupo finito, cuyo orden es un número primo, es cíclico.* En efecto, este grupo tiene que coincidir con el subgrupo cíclico engendrado por cualquiera de sus elementos, diferente de la unidad. En virtud de la descripción obtenida anteriormente de los grupos cíclicos, resulta que, *para cualquier número primo p, existe solamente un grupo finito de orden p, salvo un isomorfismo.*

§ 65. Divisores normales, grupo cociente, homomorfismos

Un subgrupo A de un grupo G se llama *divisor normal* de este grupo (o *subgrupo invariante**), si la descomposición del grupo G en clases a la izquierda respecto del subgrupo A coincide con la descomposición correspondiente a la derecha.

Por lo tanto, todos los subgrupos de un grupo abeliano son divisores normales del mismo. Por otra parte, en cualquier grupo G, el subgrupo unidad y el grupo mismo son divisores normales: ambas descomposiciones del grupo G en clases respecto del subgrupo unidad coinciden con la descomposición del grupo en elementos separados, ambas descomposiciones del grupo G en clases con respecto de este mismo grupo constan de una sola clase G.

Señalemos unos ejemplos más interesantes de divisores normales en grupos no conmutativos. En el grupo simétrico de 3er grado S_3, el subgrupo cíclico del elemento (123), que consta de la sustitución idéntica y de las sustituciones (123) y (132), representa un divisor normal: en ambas descomposiciones del grupo S_3 en clases con respecto de este subgrupo, la segunda clase adjunta consta de las sustituciones (12), (13) y (23).

En general, en el grupo simétrico S_n de grado n, el grupo alternado A_n de grado n es un divisor normal. En efecto, el orden del grupo A_n es igual a $\frac{1}{2}n!$, por lo cual, cada clase adjunta del subgrupo A_n en el grupo S_n tiene que estar constituida de la misma cantidad de elementos y, por consiguiente, solamente existe una clase más de éstas, que es precisamente el conjunto de las sustituciones impares.

En el grupo multiplicativo de las matrices cuadradas no degeneradas de orden n, cuyos elementos pertenecen al cuerpo P, las matri-

---

* También se llama *subgrupo normal*, o *subgrupo distinguido*. (*Nota del T.*).

ces, cuyos determinantes son iguales a 1, forman, evidentemente, un subgrupo. Este es, incluso, un divisor normal, puesto que las clases adjuntas a la derecha y a la izquierda de este subgrupo, engendradas por la matriz M, representan, tanto una como otra, la clase de todas las matrices, cuyos determinantes son iguales al determinante de la matriz M: es suficiente recordar que al multiplicar las matrices se multiplican sus determinantes.

A la definición de divisor normal expuesta anteriormente se le puede dar la forma siguiente:

Un subgrupo A de un grupo G se llama divisor normal de este grupo, si para cada elemento x de G

$$xA = Ax, \tag{1}$$

es decir, que para cada elemento x de G y para cada elemento a de A, se pueden elegir en A unos elementos a' y a'' tales que

$$xa = a'x, \quad ax = xa''. \tag{2}$$

Se pueden indicar también otras definiciones de divisor normal, equivalentes a la inicial. Así, llamaremos *conjugados* a los elementos a y b del grupo G, si existe en G al menos un elemento x tal, que

$$b = x^{-1}ax; \tag{3}$$

suele decirse que b es el elemento *transformado* del elemento a mediante (o por) el elemento x. Es evidente, que de (3) se deduce la igualdad

$$a = xbx^{-1} = (x^{-1})^{-1}\, bx^{-1}.$$

Un subgrupo A del grupo G es un divisor normal de éste cuando, y sólo cuando, junto con cada uno de sus elementos a contiene también a todos los elementos conjugados del mismo en G.

En efecto, si A es un divisor normal en G, entonces, en virtud de (2), para un elemento elegido a de A y para cualquier elemento x de G, se puede hallar en A un elemento a'' tal, que

$$ax = xa''.$$

De aquí que

$$x^{-1}ax = a'',$$

es decir, que cada elemento conjugado con a pertenece a A. Recíprocamente, si el subgrupo A, junto con cada uno de sus elementos a, contiene también todos los elementos conjugados con él, entonces, A contiene, en particular, al elemento

$$x^{-1}ax = a'',$$

de donde se deduce la segunda de las igualdades (2). Por la misma causa, A contiene también al elemento

$$(x^{-1})^{-1}\, ax^{-1} = xax^{-1} = a',$$

de donde se deduce la primera de las igualdades (2).

Aplicando este resultado, es fácil demostrar que *la intersección de cualesquiera divisores normales del grupo G también es un divisor normal de este grupo.* En efecto, si A y B son divisores normales del grupo G, entonces, como se ha mostrado en el párrafo anterior, la intersección $A \cap B$ representa un subgrupo del grupo G. Sea c un elemento cualquiera de $A \cap B$ y sea x un elemento cualquiera del grupo G. Entonces, el elemento $x^{-1}cx$ tiene que pertenecer tanto a A como a B, puesto que ambos divisores normales contienen al elemento c. De aquí se deduce, que el elemento $x^{-1}cx$ pertenece a la intersección $A \cap B$.

Grupo cociente (o grupo factor) *. La importancia del concepto de divisor normal se debe a que, de un modo muy natural, con las clases adjuntas relativas a un divisor normal (en virtud de (1), se puede no hacer distinción entre las clases adjuntas a la izquierda y a la derecha), se puede formar un nuevo grupo.

Obsérvese primero, que si A es un subgrupo arbitrario de un grupo G, se tiene,

$$AA = A, \tag{4}$$

pues, el producto de dos elementos cualesquiera del subgrupo A pertenece a A y, por otra parte, multiplicando todos los elementos de A por la unidad, se obtiene ya todo el subgrupo A.

Supongamos ahora que A sea un divisor normal del grupo G. *En este caso, el producto de dos clases adjuntas cualesquiera, relativas al subgrupo A* (en el sentido de multiplicación de subconjuntos del grupo G), *representa también una clase adjunta respecto de A*. En efecto, aplicando la ley asociativa del producto de subconjuntos del grupo, la igualdad (4) y la igualdad

$$yA = Ay$$

(compárese con (1)), entonces, para cualesquiera elementos x e y del grupo G, obtenemos:

$$xA \cdot yA = xyAA = xyA. \tag{5}$$

La igualdad (5) muestra que, para hallar el producto de dos clases adjuntas dadas del divisor normal A en el grupo G, se deben elegir en estas clases sendos **representantes** de un modo arbitrario (recordemos, que toda clase adjunta es engendrada por uno cualquiera de sus elementos) y se debe tomar la clase que contenga al producto de estos representantes.

* A pesar de que el autor emplea solamente la denominación de grupo factor, sin embargo, a continuación, utilizaremos la denominación de grupo cociente, que es más corriente en castellano y que, por cierto, designa lo mismo. Véase la versión castellana de la obra de Birkhoff y MacLane «Algebra Moderna», traducida por R. Rodríguez Vidal, Editorial Teide, Barcelona, pág. 171. (*Nota del. T.*).

De este modo, en el conjunto de todas las clases adjuntas del divisor normal A en el grupo G, se ha definido una operación de multiplicar. Demostremos que, *en este caso, se cumplen todas las condiciones inherentes a la definición de grupo.* En efecto, la asociatividad de la multiplicación de las clases adjuntas se deduce de la asociatividad de la multiplicación de los subconjuntos del grupo. El papel de la unidad lo desempeña el mismo divisor normal A, que representa una clase adjunta en la descomposición de G respecto de A: precisamente, en virtud de (4) y (1), para cualquier x de G, se tiene:

$$xA \cdot A = xA, \quad A \cdot xA = xAA = xA.$$

Finalmente, el inverso para la clase adjunta xA es la clase adjunta $x^{-1}A$, pues,

$$xA \cdot x^{-1}A = 1 \cdot A = A.$$

El grupo que hemos formado se denomina *grupo cociente* del grupo G por el divisor normal A y se designa con la notación G/A.

Vemos, pues, que con cada grupo se asocia toda una serie de grupos nuevos: sus grupos cocientes por diversos divisores normales. Es comprensible que, en este caso, el grupo cociente del grupo G por el subgrupo unidad es isomorfo al mismo grupo G.

Todo grupo cociente G/A de un grupo abeliano G es también abeliano, puesto que de $xy = yx$ se deduce que

$$xA \cdot yA = xyA = yxA = yA \cdot xA.$$

Todo grupo cociente G/A de un grupo cíclico G es también cíclico, puesto que si G es engendrado por el elemento g, $G = \{g\}$, y si se ha dado una clase adjunta arbitraria xA, existe un número entero k tal, que

$$x = g^k$$

y por eso,

$$xA = (gA)^k.$$

El orden de cualquier grupo cociente G/A de un grupo finito G es un divisor del orden del grupo mismo. En efecto, el orden del grupo cociente G/A es igual al índice del divisor normal A en el grupo G y, por eso, se puede aplicar la igualdad (7) del párrafo anterior.

Veamos unos cuantos **ejemplos de grupos cocientes.** Como en el grupo aditivo de los números enteros, el subgrupo de los números que son múltiplos de un número natural k tiene el índice k (véase el párrafo anterior), el grupo cociente de nuestro grupo por este subgrupo es un grupo finito de orden k que, además, es cíclico, puesto que el mismo grupo considerado es cíclico.

El grupo cociente del grupo simétrico S_n de grado n por el grupo *alternado A_n de grado n, es un grupo de 2º orden, que, como el número* 2 es primo, representa un grupo cíclico (véase el final del párrafo precedente).

Anteriormente habían sido descritas las clases adjuntas en el grupo multiplicativo de las matrices no degeneradas de orden n, cuyos elementos pertenecen a un campo P, relativas al divisor normal formado por las matrices cuyos determinantes son iguales a 1. De esta descripción se deduce que el grupo cociente correspondiente es isomorfo al grupo multiplicativo de los números del campo P que son diferentes de cero.

Homomorfismos. Los conceptos de divisor normal y de grupo cociente están estrechamente ligados con la siguiente generalización del concepto de isomorfismo.

Una *aplicación φ de un grupo G sobre un grupo G', que hace corresponder* a cada elemento a de G un elemento unívocamente determinado $a' = a\varphi$ de G', se llama *homomorfismo de G sobre G'*, si en esta aplicación cada elemento a' de G' sirve de imagen de cierto elemento a de G, $a' = a\varphi$, y si, para cualesquiera elementos a, b del grupo G,

$$(ab)\varphi = a\varphi \cdot b\varphi.$$

Es obvio, que si se requiriese además que la aplicación φ fuese biyectiva, obtendríamos la definición ya conocida de isomorfismo.

Si φ es un homomorfismo del grupo G sobre el grupo G' y 1 y a son, respectivamente, la unidad y un elemento arbitrario del grupo G, siendo $1'$ la unidad del grupo G', se tiene:

$$1\varphi = 1',$$

$$(a^{-1})\varphi = (a\varphi)^{-1}.$$

En efecto, si $1\varphi = e'$ y x' es un elemento arbitrario del grupo G', entonces existe en G un elemento x tal, que $x\varphi = x'$. De aquí que

$$x' = x\varphi = (x \cdot 1)\varphi = x\varphi \cdot 1\varphi = x' \cdot e'.$$

De un modo análogo

$$x' = e'x'$$

y, por consiguiente, $e' = 1'$.

Por otra parte, si $(a^{-1})\varphi = b'$, se tiene

$$1' = 1\varphi = (aa^{-1})\varphi = a\varphi \cdot (a^{-1})\varphi = a\varphi \cdot b'$$

y, de un modo análogo,

$$1' = b' \cdot a\varphi,$$

de donde $b' = (a\varphi)^{-1}$.

Llamemos *núcleo* del homomorfismo φ del grupo G sobre el grupo G' al conjunto de los elementos del grupo G, a los que en la aplicación φ corresponde la unidad $1'$ del grupo G'.

El núcleo de cualquier homomorfismo φ del grupo G es un divisor normal del grupo G.

En efecto, si los elementos a, b del grupo G pertenecen al núcleo del homomorfismo φ, o sea,

$$a\varphi = b\varphi = 1',$$

se tiene

$$(ab)\varphi = a\varphi \cdot b\varphi = 1' \cdot 1' = 1',$$

es decir, el producto ab también pertenece al núcleo del homomorfismo φ. Por otra parte, si $a\varphi = 1'$, se tiene

$$(a^{-1})\varphi = (a\varphi)^{-1} = 1'^{-1} = 1',$$

es decir, a^{-1} pertenece al núcleo del homomorfismo φ. Por fin, si $a\varphi = 1'$ y x es un elemento arbitrario del grupo G, entonces

$$(x^{-1}ax)\varphi = (x^{-1})\varphi \cdot a\varphi \cdot x\varphi = (x\varphi)^{-1} \cdot 1' \cdot x\varphi = 1'.$$

En resumen, tenemos que el núcleo del homomorfismo considerado representa un subgrupo del grupo G que, junto con cada uno de sus elementos, contiene también a los elementos conjugados; es, pues, un divisor normal.

Sea, ahora, A un divisor normal arbitrario del grupo G. Haciendo corresponder a cada elemento x del grupo G la clase adjunta xA, relativa al divisor normal A, a la que pertenece el mismo elemento, obtenemos una aplicación del grupo G sobre todo el grupo cociente G/A. De la definición de la multiplicación en el grupo G/A (véase (5)), se deduce que esta aplicación es un homomorfismo.

El homomorfismo obtenido se llama *homomorfismo natural* del grupo G sobre el grupo cociente G/A. Es evidente, que el mismo divisor normal A sirve de núcleo de este homomorfismo.

De aquí que *los divisores normales del grupo G, y sólo ellos, sirven de núcleos de homomorfismos de este grupo.* Este resultado se puede considerar como una definición más de divisor normal.

Resulta que con los grupos cocientes del grupo G se agotan todos los grupos sobre los que puede aplicarse el grupo G de un modo homomorfo, y con los homomorfismos naturales sobre sus grupos cocientes se agotan todos los homomorfismos del mismo. Precisando, se verifica el siguiente

Teorema de los homomorfismos. *Supongamos que se haya dado un homomorfismo φ del grupo G sobre el grupo G' y que A es el núcleo de este homomorfismo. Entonces, el grupo G' es isomorfo al grupo cociente G/A, y además, existe una aplicación isomorfa σ del primero de estos grupos sobre el segundo tal, que el resultado de la realización consecutiva de las aplicaciones φ y σ coincide con el homomorfismo natural del grupo G sobre el grupo cociente G/A.*

1/227—252

En efecto, sea x' un elemento arbitrario del grupo G', y x, un elemento tal del grupo G, que $x\varphi = x'$. Como para cualquier elemento a del núcleo A del homomorfismo φ se verifica la igualdad $a\varphi = 1'$, se tiene

$$(xa)\,\varphi = x\varphi \cdot a\varphi = x' \cdot 1' = x',$$

o sea, que todos los elementos de la clase adjunta xA se representan en φ por el elemento x'.

Por otra parte, si z es un elemento cualquiera del grupo G tal, que $z\varphi = x'$, se tiene

$$(x^{-1}z)\,\varphi = x^{-1}\varphi \cdot z\varphi = (x\varphi)^{-1} \cdot z\varphi = x'^{-1} \cdot x' = 1',$$

o sea, que $x^{-1}z$ pertenece al núcleo A del homomorfismo φ. Poniendo $x^{-1}z = a$, se tiene $z = xa$, o sea, el elemento z pertenece a la clase adjunta xA. Por consiguiente, reuniendo todos los elementos del grupo G que en el homomorfismo φ se transforman en un elemento fijado x' del grupo G', obtenemos exactamente la clase adjunta xA.

La correspondencia σ que asocia a cada elemento x' de G' la clase adjunta del grupo G relativa al divisor normal A, que consta de todos los elementos del grupo G, que en la aplicación φ tienen por imagen a x', es una aplicación biyectiva del grupo G' sobre el grupo G/A. Esta aplicación σ es un isomorfismo, puesto que si

$$x'\sigma = xA, \quad y'\sigma = yA,$$

o sea, si

$$x\varphi = x', \quad y\varphi = y',$$

entonces,

$$(xy)\,\varphi = x\varphi \cdot y\varphi = x'y',$$
$$(x'y')\,\sigma = xyA = xA \cdot yA = x'\sigma \cdot y'\sigma.$$

Finalmente, si x es un elemento arbitrario de G y $x\varphi = x'$, se tiene

$$(x\varphi)\,\sigma = x'\sigma = xA,$$

es decir, que en la realidad, la realización consecutiva del homomorfismo φ y del isomorfismo σ hace corresponder al elemento x la clase adjunta xA engendrada por él mismo. El teorema queda demostrado.

§ 66. Sumas directas de grupos abelianos

Queremos acabar este capítulo con un teorema de la teoría de los grupos más profundo que aquellas propiedades elementales de los grupos que se habían expuesto anteriormente. A saber, basándose en la descripción de los grupos cíclicos, ya conocida por el § 64,

obtendremos en el párrafo siguiente una *descripción completa de los grupos finitos abelianos.*

Como está convenido en la teoría de los grupos abelianos, para la operación en el grupo se empleará la forma de expresión aditiva: se hablará de la **suma** $a+b$ de los elementos a y b del grupo, del **subgrupo nulo** 0, de los **múltiplos** ka de cierto elemento a, etc, etc.

En este párrafo examinaremos una construcción, cuya exposición va a estar adaptada para los grupos abelianos, a pesar de que podría ser presentada a la vez para grupos cualesquiera (aunque no fuesen conmutativos). Esta construcción está dictada por los ejemplos que siguen. El plano, considerado como un espacio lineal real de dos dimensiones, representa un grupo abeliano respecto a la suma de vectores. En este plano, cualquier recta que pase por el origen de coordenadas es un subgrupo del grupo indicado. Si A_1 y A_2 son dos rectas de éstas, entonces, como se sabe, todo vector que parte del origen de coordenadas se representa unívocamente en forma de suma de sus proyecciones sobre las rectas A_1 y A_2. Análogamente, todo vector del espacio lineal de tres dimensiones se expresa unívocamente en forma de una suma de tres vectores que pertenecen a tres rectas dadas A_1, A_2, A_3, suponiendo que estas rectas no estén situadas en un plano.

Se dice que un grupo abeliano G es una *suma directa* de sus subgrupos $A_1, A_2, \ldots, A_k$,

$$G=A_1+A_2+\ldots+A_k, \tag{1}$$

si cada elemento x del grupo G se expresa, y además, **unívocamente**, en forma de una suma de elementos $a_1, a_2, \ldots, a_k$ tomados en los subgrupos $A_1, A_2, \ldots, A_k$, correspondientemente:

$$x=a_1+a_2+\ldots+a_k. \tag{2}$$

La expresión (1) se denomina *descomposición directa* del grupo G; los subgrupos A_i, $i=1, 2, \ldots, k$, se llaman *sumandos directos* de esta descomposición, y el elemento a_i de (2), *componente* del elemento x en el sumando directo A_i de la descomposición (1), $i=1, 2, \ldots, k$.

Si se ha dado una descomposición directa (1) *del grupo* G, *y si todos, o unos cuantos, sumandos directos* A_i *de esta descomposición están también descompuestos en una suma directa*

$$A_i=A_{i1}+A_{i2}+\ldots+A_{ik_i}, \quad k_i \geqslant 1, \tag{3}$$

entonces, el grupo G *representa una suma directa de todos sus subgrupos*

$$A_{ij},\ j=1, 2, \ldots, k_i,\ i=1, 2, \ldots, k.$$

En efecto, para un elemento arbitrario x del grupo G existe una expresión (2) respecto a la descomposición directa (1), y para cada

En efecto, sea x' un elemento arbitrario del grupo G', y x, un elemento tal del grupo G, que $x\varphi = x'$. Como para cualquier elemento a del núcleo A del homomorfismo φ se verifica la igualdad $a\varphi = 1'$, se tiene

$$(xa)\,\varphi = x\varphi \cdot a\varphi = x' \cdot 1' = x',$$

o sea, que todos los elementos de la clase adjunta xA se representan en φ por el elemento x'.

Por otra parte, si z es un elemento cualquiera del grupo G tal, que $z\varphi = x'$, se tiene

$$(x^{-1}z)\,\varphi = x^{-1}\varphi \cdot z\varphi = (x\varphi)^{-1} \cdot z\varphi = x'^{-1} \cdot x' = 1',$$

o sea, que $x^{-1}z$ pertenece al núcleo A del homomorfismo φ. Poniendo $x^{-1}z = a$, se tiene $z = xa$, o sea, el elemento z pertenece a la clase adjunta xA. Por consiguiente, reuniendo todos los elementos del grupo G que en el homomorfismo φ se transforman en un elemento fijado x' del grupo G', obtenemos exactamente la clase adjunta xA.

La correspondencia σ que asocia a cada elemento x' de G' la clase adjunta del grupo G relativa al divisor normal A, que consta de todos los elementos del grupo G, que en la aplicación φ tienen por imagen a x', es una aplicación biyectiva del grupo G' sobre el grupo G/A. Esta aplicación σ es un isomorfismo, puesto que si

$$x'\sigma = xA, \quad y'\sigma = yA,$$

o sea, si

$$x\varphi = x', \quad y\varphi = y',$$

entonces,

$$(xy)\,\varphi = x\varphi \cdot y\varphi = x'y',$$

$$(x'y')\,\sigma = xyA = xA \cdot yA = x'\sigma \cdot y'\sigma.$$

Finalmente, si x es un elemento arbitrario de G y $x\varphi = x'$, se tiene

$$(x\varphi)\,\sigma = x'\sigma = xA,$$

es decir, que en la realidad, la realización consecutiva del homomorfismo φ y del isomorfismo σ hace corresponder al elemento x la clase adjunta xA engendrada por él mismo. El teorema queda demostrado.

§ 66. Sumas directas de grupos abelianos

Queremos acabar este capítulo con un teorema de la teoría de los grupos más profundo que aquellas propiedades elementales de los grupos que se habían expuesto anteriormente. A saber, basándose en la descripción de los grupos cíclicos, ya conocida por el § 64,

Un grupo abeliano G representa una suma directa de sus subgrupos $A_1, A_2, \ldots, A_k$ cuando, y sólo cuando, el mismo es engendrado por estos subgrupos,

$$G = \{A_1, A_2, \ldots, A_k\}, \tag{6}$$

y la intersección de cada subgrupo A_i, $i = 2, \ldots, k$, con el subgrupo engendrado por todos los subgrupos anteriores $A_1, A_2, \ldots, A_{i-1}$, contiene solamente al cero,

$$\{A_1, A_2, \ldots, A_{i-1}\} \cap A_i = 0, \ i = 2, \ldots, k. \tag{7}$$

En efecto, si el grupo G posee una descomposición directa (1), entonces, para cada elemento x de G existe una expresión (2) y, por esto, se verifica la igualdad (6). El cumplimiento de la igualdad (7) es consecuencia de la unicidad de la expresión (2) para cualquier elemento x: si para cierto i, la intersección $\{A_1, A_2, \ldots, A_{i-1}\} \cap \cap A_i$ contuviese un elemento x no nulo, entonces, por una parte, x se podría expresar como un elemento a_i de A_i, o sea, $x = a_i$, y por eso,

$$x = 0 + \ldots + 0 + a_i + 0 + \ldots + 0; \tag{8}$$

por otra parte, x, como elemento del subgrupo $\{A_1, A_2, \ldots, A_{i-1}\}$ posee una expresión de la forma

$$x = a_1 + a_2 + \ldots + a_{i-1},$$

o sea,

$$x = a_1 + a_2 + \ldots + a_{i-1} + 0 + \ldots + 0. \tag{9}$$

Es evidente, que para el elemento x, (8) y (9) son dos expresiones distintas de la forma (2).

Recíprocamente, supongamos que se cumplen las igualdades (6) y (7). De (6) se deduce, que cualquier elemento x del grupo G posee por lo menos una expresión de la forma (2). Por otra parte, supongamos que para cierto elemento x existen dos expresiones distintas de la forma (2)

$$x = a_1 + a_2 + \ldots + a_k = a_1' + a_2' + \ldots + a_k'. \tag{10}$$

Entonces, se puede hallar tal i, $i \leqslant k$, que

$$a_k = a_k', \ a_{k-1} = a_{k-1}', \ \ldots, a_{i+1} = a_{i+1}', \tag{11}$$

pero

$$a_i \neq a_i',$$

o sea,

$$a_i - a_i' \neq 0. \tag{12}$$

Sin embargo, de (9) y (11) se deduce la igualdad

$$a_i - a_i' = (a_1' - a_1) + (a_2' - a_2) + \ldots + (a_{i-1}' - a_{i-1}).$$

que, en virtud de (12), contradice a la igualdad (7). El teorema queda demostrado.

El concepto de suma directa se puede examinar de otro modo distinto. Sean dados k grupos abelianos arbitrarios $A_1, A_2, \ldots, A_k$, algunos de los cuales pueden ser isomorfos. Designemos con G el conjunto de todos los sistemas posibles de la forma

$$(a_1, a_2, \ldots, a_k), \tag{13}$$

formados por sendos elementos de los grupos $A_1, A_2, \ldots, A_k$. El conjunto G se convierte en un grupo abeliano, si la suma de los sistemas de la forma (13) se define por la regla:

$$(a_1, a_2, \ldots, a_k) + (a'_1, a'_2, \ldots, a'_k) = \\ = (a_1 + a'_1, a_2 + a'_2, \ldots, a_k + a'_k), \tag{14}$$

según la cual se suman los elementos de los grupos dados $A_1, A_2, \ldots$ $\ldots, A_k$ por separado. En efecto, las leyes asociativa y conmutativa de esta suma se deducen del cumplimiento de estas leyes en cada uno de los grupos dados; el papel del cero lo desempeña el sistema

$$(0_1, 0_2, \ldots, 0_k),$$

donde mediante 0_i se señala el elemento nulo del grupo A_i, $i = 1, 2, \ldots, k$; el elemento opuesto para el sistema (13) es el sistema

$$(-a_1, -a_2, \ldots, -a_k).$$

El grupo abeliano G construido se llama *suma directa* de los grupos $A_1, A_2, \ldots, A_k$ y se designa, como anteriormente, mediante

$$G = A_1 + A_2 + \ldots + A_k.$$

La razón de esta denominación consiste en que *el grupo G, que representa una suma directa de los grupos $A_1, A_2, \ldots, A_k$ en el sentido que acabamos de definir, se puede descomponer en una suma directa de sus subgrupos $A'_1, A'_2, \ldots, A'_k$, que son isomorfos a los grupos $A_1, A_2, \ldots, A_k$, correspondientemente.*

Designemos, para esto, mediante A'_i, $i = 1, 2, \ldots, k$, el conjunto de los elementos del grupo G, o sea, de los sistemas de la forma (13), en los que en el lugar de i figura un elemento arbitrario a_i del grupo A_i, y en los demás lugares, los ceros de los grupos correspondientes; éstos son, por consiguiente, los sistemas de la forma

$$(0_1, \ldots, 0_{i-1}, a_i, 0_{i+1}, \ldots, 0_k). \tag{15}$$

La definición de la suma (14) muestra que el conjunto A'_i representa un subgrupo del grupo G; el isomorfismo de este subgrupo con el grupo A_i se obtiene haciendo corresponder a cada sistema (15) el elemento a_i del grupo A_i.

Queda por demostrar que el grupo G representa una suma directa de los subgrupos A_1', A_2', ..., A_k'. En efecto, cualquier elemento (13) del grupo G se puede representar en forma de una suma de elementos de los subgrupos indicados:

$$(a_1, a_2, \ldots, a_k) = (a_1, 0_2, \ldots, 0_k) + \\ + (0_1, a_2, 0_3, \ldots, 0_k) + \ldots + (0_1, 0_2, \ldots, 0_{k-1}, a_k).$$

La unicidad de esta representación se deduce de que diferentes sistemas de la forma (13) son **diferentes** elementos del grupo G.

Si se han dado dos sistemas de grupos abelianos, A_1, A_2, ..., A_k *y* B_1, B_2, ..., B_k, *y los grupos* A_i *y* B_i, $i = 1, 2, \ldots, k$, *son isomorfos, entonces los grupos*

$$G = A_1 + A_2 + \ldots + A_k$$

y

$$H = B_1 + B_2 + \ldots + B_k$$

también son isomorfos.

En efecto, si para $i = 1, 2, \ldots, k$, se ha establecido un isomorfismo φ_i entre los grupos A_i y B_i que hace corresponder a cada elemento a_i de A_i el elemento $a_i\varphi_i$ de B_i, entonces es evidente, que la aplicación φ que a cada elemento $(a_1, a_2, \ldots, a_k)$ del grupo G asocia el elemento del grupo H determinado por la igualdad

$$(a_1, a_2, \ldots, a_k)\varphi = (a_1\varphi_1, a_2\varphi_2, \ldots, a_k\varphi_k),$$

es un isomorfismo que aplica al grupo G sobre el grupo H.

Si se han dado los grupos abelianos **finitos** A_1, A_2, ..., A_k, *cuyos órdenes correspondientes son* n_1, n_2, ..., n_k, *entonces la suma directa G de estos grupos es también un grupo finito y su orden n es igual al producto de los órdenes de los sumandos directos,*

$$n = n_1 n_2 \ldots n_k. \tag{16}$$

En efecto, el número de sistemas diversos de la forma (13), para cada uno de los cuales el elemento a_1 puede tomar n_1 valores distintos, el elemento a_2, toma n_2 valores distintos, etc. se determina por la igualdad (16).

Veamos unos cuantos **ejemplos.**

Si el orden n de un grupo cíclico finito $\{a\}$ se descompone en un producto de dos números naturales que son primos entre sí,

$$n = st, \ (s, t) = 1,$$

entonces, el grupo $\{a\}$ se descompone en una suma directa de dos grupos cíclicos, cuyos órdenes correspondientes son s y t.

Para el grupo $\{a\}$ emplearemos la expresión aditiva. Poniendo $b = ta$, se tiene

$$sb = (st)a = na = 0,$$

pero, para $0 < k < s$,

$$kb = (kt)\, a \neq 0,$$

es decir, el subgrupo cíclico $\{b\}$ tiene el orden s. Análogamente, el subgrupo cíclico $\{c\}$ del elemento $c = sa$ tiene el orden t. La intersección $\{b\} \cap \{c\}$ contiene sólo el cero, puesto que si $kb = lc$ para $0 < k < s$, $0 < l < t$, entonces

$$(kt)\, a = (ls)\, a,$$

y como los números kt y ls son menores que n, se tiene

$$kt = ls,$$

lo cual es imposible, ya que los números s y t son primos entre sí. Finalmente, existen unos números u y v tales, que

$$su + tv = 1,$$

y, por lo tanto,

$$a = v\,(ta) + u\,(sa) = vb + uc,$$

y, por consiguiente, cualquier elemento del grupo $\{a\}$ se puede representar como una suma de elementos de los subgrupos $\{b\}$ y $\{c\}$.

Llamaremos a un grupo abeliano G *indescomponible*, si no puede ser descompuesto en una suma directa de dos o de unos cuantos subgrupos, diferentes del subgrupo cero. Un grupo cíclico finito, cuyo orden es una potencia de un número primo p, se denomina grupo cíclico *primario* respecto al número primo p. Aplicando unas cuantas veces la proposición demostrada anteriormente, obtenemos, que *todo grupo cíclico finito se descompone en una suma directa de grupos cíclicos primarios, respecto a diversos números primos.* Más exactamente, *todo grupo cíclico de orden*

$$n = p_1^{k_1} p_2^{k_2} \ldots p_s^{k_s},$$

donde $p_1, p_2, \ldots, p_s$ son números primos **distintos**, *se descompone en una suma directa de s grupos cíclicos que tienen los órdenes $p_1^{k_1}, p_2^{k_2}, \ldots, p_s^{k_s}$, respectivamente.*

Todo grupo cíclico primario es indescomponible.

En efecto, sea dado un grupo cíclico finito $\{a\}$ de orden p^k, donde p es un número primo. Si este grupo fuese descomponible, entonces, en virtud de (7), tendría subgrupos diferentes de cero, la intersección de los cuales sería igual a cero. Sin embargo, en la realidad, **todo subgrupo diferente de cero contiene el elemento diferente de cero**

$$b = p^{k-1} a.$$

Para la demostración, tomemos un elemento arbitrario x diferente de cero de nuestro grupo,

$$x = sa, \ 0 < s < p^k.$$

El número s se puede escribir en la forma

$$s = p^l s', \ 0 \leqslant l < k,$$

donde el número s' ya no es divisible por p y, por consiguiente, éstos son primos entre sí, debido a lo cual, existen unos números u y v tales, que

$$s'u + pv = 1.$$

Entonces,

$$(p^{k-l-1}u)\, x = (p^{k-l-1}us)\, a = (p^{k-1}us')\, a =$$
$$= p^{k-1}(1 - pv)\, a = (p^{k-1} - p^k v)\, a = p^{k-1}a - v\,(p^k a) = p^{k-1}a = b,$$

o sea, el elemento b pertenece al subgrupo cíclico $\{x\}$.

El grupo aditivo de los números enteros (o sea, el grupo cíclico infinito), y también el grupo aditivo de todos los números racionales, son grupos indescomponibles.

Esto se deduce de que en cada uno de estos grupos, para cualquier par de elementos diferentes de cero, existe un común múltiplo diferente de cero, es decir, dos subgrupos cíclicos cualesquiera, diferentes de cero, tienen una intersección diferente de cero.

Obsérvese que, si en el grupo abeliano G la operación se llama multiplicación, entonces, se debe hablar del *producto directo* y no de la suma directa.

El grupo multiplicativo de los números reales diferentes de cero se descompone en un producto directo del grupo multiplicativo de los números reales positivos y del grupo formado por los números 1 y −1, respecto a la multiplicación.

En efecto, a la intersección de los dos subgrupos indicados de nuestro grupo pertenece solamente el número 1, que es el elemento unidad de este grupo. Por otra parte, todo número positivo es igual al producto de sí mismo por el número 1, todo número negativo es igual al producto de su valor absoluto por el número −1.

§ 67. Grupos abelianos finitos

Tomando cualquier conjunto finito de grupos cíclicos primarios, algunos de los cuales pueden estar referidos a un mismo número primo, o incluso pueden tener un mismo orden, o sea, que pueden ser isomorfos, la suma directa de ellos representa un grupo abeliano finito. Resulta, que con esto se agotan todos los grupos abelianos finitos:

Teorema fundamental de los grupos abelianos finitos. *Todo grupo abeliano finito G que no es un grupo cero, se descompone en una suma directa de subgrupos cíclicos primarios.*

Comenzaremos la demostración de este teorema observando que *en el grupo G, indispensablemente, existen elementos diferentes de cero, cuyos órdenes son potencias de números primos.* En efecto, si un elemento x del grupo G, diferente de cero, tiene el orden l, $lx = 0$, y p^k, $k > 0$, es una potencia del número primero p tal, que el número l,

$$l = p^k m,$$

es divisible por ella, entonces, el elemento mx es diferente de cero y tiene el orden p^k.

Sean

$$p_1, \; p_2, \; \ldots, \; p_s \tag{1}$$

todos los números primos **diversos**, algunas de cuyas potencias sirven de órdenes de algunos elementos del grupo G. Designemos con p cualquiera de estos números, y con P, el conjunto de los elementos del grupo G, cuyos órdenes son potencias del número p.

El conjunto P representa un subgrupo del grupo G. En efecto, P contiene al elemento 0, ya que su orden es igual a $1 = p^0$. Por otra parte, si $p^k x = 0$, entonces, $p^k(-x) = 0$. Finalmente, si $p^k x = 0$, $p^l y = 0$, y si, por ejemplo, $k \geqslant l$, entonces,

$$p^k (x + y) = 0,$$

o sea, el orden del elemento $x + y$, o bien es el número p^k, o bien es un divisor de este número, es decir, es una potencia del número p.

Tomando, por p cada uno de los números (1), sucesivamente, obtenemos s subgrupos no nulos,

$$P_1, \; P_2, \; \ldots, \; P_s. \tag{2}$$

El grupo G es una suma directa de estos subgrupos,

$$G = P_1 + P_2 + \ldots + P_s \tag{3}$$

En efecto, si x es un elemento arbitrario del grupo G, su orden l sólo puede dividirse por ciertos números primos del sistema (1),

$$l = p_1^{k_1} p_2^{k_2} \ldots p_s^{k_s},$$

donde $k_i \geqslant 0$, $i = 1, 2, \ldots, s$. Por eso, como se había demostrado al final del párrafo anterior, el subgrupo cíclico $\{x\}$ se descompone en una suma directa de subgrupos cíclicos primarios que tienen los órdenes $p_1^{k_1}$, $p_2^{k_2}$, $\ldots$, $p_s^{k_s}$ respectivamente. Estos subgrupos cíclicos primarios pertenecen a los subgrupos (2) correspondientes y, por

consiguiente, el elemento x se representa en forma de una suma de elementos, tomados uno por uno en todos o en unos cuantos de los subgrupos (2). De este modo, queda demostrada la igualdad

$$G=\{P_1, P_2, \ldots, P_s\},$$

que es análoga a la igualdad (6) del párrafo anterior.

Para demostrar la igualdad, análoga a la igualdad (7) del mismo párrafo, tomemos cualquier i, $2 \leqslant i \leqslant s$. Entonces, cualquier elemento y del subgrupo $\{P_1, P_2, \ldots, P_{i-1}\}$ tiene la forma

$$y=a_1+a_2+\ldots+a_{i-1},$$

donde el elemento a_j, $j=1, 2, \ldots, i-1$, pertenece al subgrupo P_j, es decir, tiene el orden $p_j^{k_j}$. Entonces,

$$(p_1^{k_1} p_2^{k_2} \ldots p_{i-1}^{k_{i-1}})\, y=0,$$

o sea, el orden del elemento y es cierto divisor del número $p_1^{k_1} p_2^{k_2} \ldots$ $\ldots p_{i-1}^{k_{i-1}}$ y, por consiguiente, el elemento y, si es diferente de cero, no puede pertenecer al subgrupo P_i. De este modo, queda demostrado, que

$$\{P_1, P_2, \ldots, P_{i-1}\} \cap P_i=0,$$

que es lo que se quería demostrar.

Obsérvese que el grupo abeliano en el que los órdenes de todos los elementos son potencias de un mismo número primo p, se denomina *primario* respecto del número p. Los grupos cíclicos primarios son casos particulares de los grupos primarios. Por lo tanto, los subgrupos (2) son primarios. Estos se llaman *componentes primarios* del grupo G, y la descomposición directa (3), *descomposición de este grupo en componentes primarios.* Como los subgrupos (2) están determinados unívocamente en el grupo G, *la descomposición del grupo G en componentes primarios se determina unívocamente.*

Es comprensible, que la descomposición de todo grupo abeliano finito en una suma directa de grupos primarios **reduce la demostración del teorema fundamental al caso de un grupo abeliano finito primario** P, **respecto de cierto número primo p.** Examinemos este caso.

Sea a_1 uno de los elementos del grupo P que tienen en éste el orden máximo. Si, luego, existen en el grupo P elementos, diferentes de cero, las intersecciones de cuyos subgrupos cíclicos con el subgrupo cíclico $\{a_1\}$ son iguales a cero, entonces, mediante a_2 indicamos uno de los elementos de orden máximo entre los elementos que poseen esta propiedad; por lo tanto,

$$\{a_1\} \cap \{a_2\}=0.$$

Supongamos que ya se han elegido los elementos $a_1, a_2, \ldots, a_{i-1}$. El subgrupo del grupo P, engendrado por sus subgrupos cíclicos, lo indicaremos mediante $\{a_1, a_2, \ldots, a_{i-1}\}$,

$$\{\{a_1\}, \{a_2\}, \ldots, \{a_{i-1}\}\} = \{a_1, a_2, \ldots, a_{i-1}\}. \quad (4)$$

Es evidente, que éste se compone de todos los elementos del grupo P que se pueden expresar en forma de una suma de elementos, múltiplos de los elementos $a_1, a_2, \ldots, a_{i-1}$; diremos que este subgrupo *está engendrado* por los elementos $a_1, a_2, \ldots, a_{i-1}$. Designemos ahora con a_i uno de los elementos de orden máximo entre los elementos del grupo P, las intersecciones de cuyos subgrupos cíclicos con el subgrupo $\{a_1, a_2, \ldots, a_{i-1}\}$ son iguales a cero; por lo tanto,

$$\{a_1, a_2, \ldots, a_{i-1}\} \cap \{a_i\} = 0. \quad (5)$$

Como el grupo P es finito, este proceso tendrá fin; supongamos que esto ocurre después de que se han elegido los elementos $a_1, a_2, \ldots, a_s$. Designando con P' el subgrupo engendrado por estos elementos,

$$P' = \{a_1, a_2, \ldots, a_s\},$$

o sea,

$$P' = \{\{a_1\}, \{a_2\}, \ldots, \{a_s\}\}, \quad (6)$$

se tiene que **el subgrupo cíclico engendrado por cualquier elemento del grupo P, diferente de cero, tiene con el subgrupo P' una intersección no nula.**

En virtud de (4), la igualdad (6) y la igualdad (5), que se verifican para $i = 2, 3, \ldots, s$, muestran que **el subgrupo P' es una suma directa de los subgrupos cíclicos** $\{a_1\}, \{a_2\}, \ldots, \{a_s\}$,

$$P' = \{a_1\} \dotplus \{a_2\} \dotplus \ldots \dotplus \{a_s\}. \quad (7)$$

Queda por demostrar que **el subgrupo P' coincide en la realidad con todo el grupo P.**

Sea x un elemento cualquiera del grupo P que tenga el orden p. Como

$$P' \cap \{x\} \neq 0,$$

y el subgrupo $\{x\}$ no tiene subgrupos no nulos, diferentes de sí mismo (recordemos, que el orden de un subgrupo es divisor del orden del grupo, y que el número p es primo), el subgrupo $\{x\}$ verdaderamente está contenido en el subgrupo P' y, por consiguiente, x pertenece a P'. Por lo tanto, todos los elementos de orden p del grupo P pertenecen al subgrupo P'.

Supongamos que ya está demostrado que al subgrupo P' pertenecen todos los elementos del grupo P, cuyos órdenes no son mayores que el número p^{k-1}, y sea x un elemento cualquiera de P de orden p^k.

Como muestra la elección de los elementos $a_1, a_2, \ldots, a_s$, el orden de éstos no va creciendo y, por esto, se puede señalar tal i, $1 \leqslant i-1 \leqslant s$, que los órdenes de los elementos $a_1, a_2, \ldots, a_{i-1}$ son mayores o iguales a p^k, y para $i-1 < s$, el orden del elemento a_i es estrictamente menor que este número, es decir, es menor que el orden del elemento x. En virtud de las condiciones a que está ajustada la elección del elemento a_i, de aquí se deduce que, si

$$Q=\{a_1, a_2, \ldots, a_{i-1}\},$$

entonces,

$$Q \cap \{x\} \neq 0.$$

Sin embargo, en el párrafo anterior se había demostrado que todo subgrupo no nulo de un grupo cíclico primario $\{x\}$ de orden p^k contiene el elemento

$$y=p^{k-1}x. \tag{8}$$

Por consiguiente, este elemento y pertenece a la intersección $Q \cap \cap \{x\}$, y, por lo tanto, al subgrupo Q. Esto da la posibilidad de expresar y en forma de una suma de elementos, múltiplos de los elementos $a_1, a_2, \ldots, a_{i-1}$:

$$y=l_1a_1+l_2a_2+\ldots+l_{i-1}a_{i-1}. \tag{9}$$

De (8) se deduce, que el elemento y tiene el orden p. Por eso,

$$(pl_1)a_1+(pl_2)a_2+\ldots+(pl_{i-1})a_{i-1}=0,$$

o sea, que en virtud de la existencia de la descomposición directa (7),

$$(pl_j)a_j=0, \quad j=1, 2, \ldots, i-1.$$

Por lo tanto, el número pl_j tiene que dividirse por el orden del elemento a_j, y por esto, también por el número p^k, de donde se deduce que l_j se divide por p^{k-1},

$$l_j=p^{k-1}m_j, \quad j=1, 2, \ldots, i-1. \tag{10}$$

Sea

$$z=m_1a_1+m_2a_2+\ldots+m_{i-1}a_{i-1}.$$

Este elemento pertenece al subgrupo Q y, por consiguiente, al subgrupo P'; además, en virtud de (9) y (10),

$$y=p^{k-1}z. \tag{11}$$

De (8) y (11) se deduce la igualdad

$$p^{k-1}(x-z)=0,$$

es decir, que el orden del elemento

$$t=x-z$$

no es mayor que p^{k-1} y, por consiguiente, en virtud de la hipotesis de la inducción, t pertenece al subgrupo P'. Por esto, el elemento x, como suma de dos elementos de P', $x = z + t$, también pertenece al subgrupo P'. De este modo, queda demostrado que todos los elementos de orden p^k del grupo P pertenecen a P'. Por consiguiente, nuestra demostración por inducción da la posibilidad de afirmar que todos los elementos del grupo P pertenecen al subgrupo P', o sea, que $P' = P$. La demostración del teorema fundamental está terminada.

Como resultado complementario, obtenemos que *un grupo abeliano finito es primario respecto al número primo p, cuando, y sólo cuando, su orden es una potencia de este número p*. En efecto, se había demostrado que todo grupo abeliano finito P que es primario (respecto a p), se descompone en una suma directa de grupos cíclicos primarios (respecto a p), y por eso, el orden del grupo P es igual al producto de los órdenes de estos grupos cíclicos, o sea, es una potencia del número p. Recíprocamente, si el orden de un grupo abeliano finito es igual a p^h, donde p es un número primo, entonces, el orden de cualquiera de sus elementos es divisor de este número, es decir, también es una potencia del número p, y, por lo tanto, el grupo resulta ser primario respecto a p.

Con el teorema fundamental no se agota todavía el problema de la descripción total de los grupos abelianos finitos, puesto que todavía no se ha excluido la posibilidad de que las sumas directas de dos conjuntos distintos de grupos cíclicos, primarios respecto a ciertos números primos, sean grupos isomorfos. En la realidad esto no se verifica, como muestra el **teorema** que sigue:

Si, de dos modos distintos, se ha descompuesto un grupo abeliano finito G en una suma directa de subgrupos cíclicos primarios,

$$G = \{a_1\} + \{a_2\} + \ldots + \{a_s\} = \{b_1\} + \{b_2\} + \ldots + \{b_t\}, \qquad (12)$$

entonces, ambas descomposiciones directas poseen el mismo número de sumandos directos, $s = t$, y entre los sumandos directos de estas descomposiciones se puede establecer una correspondencia biunívoca tal, que los sumandos correspondientes sean grupos cíclicos de un mismo orden, es decir, isomorfos.

Observemos primero, que si tomamos en la primera de las descomposiciones directas (12), por ejemplo, los sumandos directos que se relacionan al número primo dado p, su suma directa será un subgrupo primario (respecto a p) del grupo G, e incluso componente primaria de este grupo, puesto que su orden es igual a la potencia máxima del número p por la que se divide el orden del grupo G. Reuniendo de este modo todos los sumandos directos en cada una de las descomposiciones (12), obtenemos en ambos casos la descomposición del grupo G en componentes primarias, cuya unicidad ya fue señalada anteriormente.

Esto nos permite demostrar el teorema, suponiendo que **el mismo grupo G es primario respecto al número primo p.** Sea elegida la numeración de los sumandos directos en cada una de las descomposiciones (12) de tal modo, que los órdenes de estos sumandos no vayan creciendo, es decir, que teniendo los elementos $a_1, a_2, \ldots, a_s$ los órdenes

$$p^{k_1},\ p^{k_2}, \ldots, p^{k_s},$$

respectivamente, sea,

$$k_1 \geqslant k_2 \geqslant \ldots \geqslant k_s,$$

y teniendo los elementos $b_1, b_2, \ldots, b_t$ los órdenes

$$p^{l_1},\ p^{l_2}, \ldots, p^{l_t},$$

respectivamente, sea,

$$l_1 \geqslant l_2 \geqslant \ldots \geqslant l_t.$$

Si no se cumpliese la tesis de nuestro teorema, se hallaría un $i \geqslant 1$, tal, que

$$k_1 = l_1, \ldots, k_{i-1} = l_{i-1}, \tag{13}$$

pero,

$$k_i \neq l_i.$$

Está claro, que $i \leqslant \min(s, t)$, puesto que para cada una de las descomposiciones (12), el producto de los órdenes de todos los *sumandos directos es igual al orden del grupo G.* Mostremos que nuestra suposición nos lleva a una contradicción.

Sea, por ejemplo,

$$k_i < l_i. \tag{14}$$

Designemos con H el conjunto de los elementos del grupo G cuyos órdenes no sobrepasan a p^{k_i}. *Este representa un subgrupo del grupo G*, puesto que si x e y son elementos de H, entonces, $x+y$ y $-x$ son de orden no superior al número p^{k_i}.

Obsérvese, que al subgrupo H pertenecen, en particular, los elementos siguientes:

$$p^{k_1-k_i}a_1,\ p^{k_2-k_i}a_2, \ldots, p^{k_{i-1}-k_i}a_{i-1}, a_i, a_{i+1}, \ldots, a_s.$$

Por otra parte, si $1 \leqslant j \leqslant i-1$, entonces, el orden del elemento $p^{k_j-k_i-1}a_j$ es igual a p^{k_i+1}, y por eso, no pertenece a H. De aquí se deduce, que la clase adjunta $a_j + H$ (¡recordemos, que estamos empleando la expresión aditiva!) tiene, como elemento del grupo cociente G/H, el orden $p^{k_j-k_i}$; este mismo orden tiene *su* subgrupo cíclico $\{a_j + H\}$. Demostremos que **el grupo G/H es una suma directa de los subgrupos cíclicos** $\{a_j + H\}$, $j = 1, 2, \ldots, i-1$,

$$G/H = \{a_1 + H\} + \{a_2 + H\} + \ldots + \{a_{i-1} + H\}, \tag{15}$$

y que, por esto, su orden es igual al número

$$p^{(k_1-k_i)+(k_2-k_i)+\ldots+(k_{i-1}-k_i)}. \tag{16}$$

Si x es un elemento arbitrario del grupo G, existe la expresión

$$x=m_1a_1+m_2a_2+\ldots+m_sa_s.$$

Supongamos que, para $j=1, 2, \ldots, i-1$,

$$m_j=p^{k_j-k_i}q_j+n_j,$$

donde

$$0\leqslant n_j<p^{k_j-k_i}. \tag{17}$$

Entonces,

$$m_ja_j=q_j(p^{k_j-k_i}a_j)+n_ja_j,$$

y como el primer sumando del segundo miembro está contenido en H, se tiene

$$m_ja_j+H=n_ja_j+H.$$

Por otra parte,

$$m_ia_i+H=H, \ldots, m_sa_s+H=H,$$

por eso,

$$x+H=(m_1a_1+H)+(m_2a_2+H)+\ldots+(m_sa_s+H)=$$
$$=(n_1a_1+H)+(n_2a_2+H)+\ldots+(n_{i-1}a_{i-1}+H). \tag{18}$$

Supongamos que existe una expresión más de éstas,

$$x+H=(n'_1a_1+H)+(n'_2a_2+H)+\ldots+(n'_{i-1}a_{i-1}+H), \tag{19}$$

donde

$$0\leqslant n'_j<p^{k_j-k_i},\ j=1,\ 2, \ldots, i-1. \tag{20}$$

Entonces, los elementos

$$n_1a_1+n_2a_2+\ldots+n_{i-1}a_{i-1}$$

y

$$n'_1a_1+n'_2a_2+\ldots+n'_{i-1}a_{i-1}$$

están en una misma clase adjunta relativa a H, o sea, su diferencia pertenece a H y, por esto,

$$p^{k_i}[(n_1-n'_1)a_1+(n_2-n'_2)a_2+\ldots+(n_{i-1}-n'_{i-1})a_{i-1}]=0.$$

De aquí se deduce (ya que la primera de las descomposiciones (12) es directa) que

$$p^{k_i}(n_j-n'_j)a_j=0,\ j=1, 2, \ldots, i-1,$$

y, por lo tanto, el número $p^{k_i}(n_j - n'_j)$ tiene que dividirse por el orden p^{k_j} del elemento a_j y, por consiguiente, la diferencia $n_j - n'_j$ se divide por el número $p^{k_j - k_i}$. En virtud de (17) y (20), de aquí se deduce que

$$n_j = n'_j, \; j = 1, 2, \ldots, i-1,$$

es decir, las expresiones (18) y (19) son idénticas. De este modo, queda demostrada la existencia de la descomposición directa (15).

Consideraciones análogas, realizadas para la segunda de las descomposiciones (12), muestran que este mismo grupo cociente G/H posee una descomposición directa

$$G/H = \{b_1 + H\} + \{b_2 + H\} + \ldots + \{b_{i-1} + H\} + \{b_i + H\} + \ldots,$$

es decir, que en virtud de (13) y (14), su orden tiene que ser **estrictamente mayor** que el número (16). Esta contradicción demuestra el teorema.

Ya hemos obtenido una exposición completa de los grupos abelianos finitos. Así, pues, *tomamos todos los conjuntos finitos posibles de números naturales*

$$(n_1, n_2, \ldots, n_k),$$

diferentes de la unidad, pero no indispensablemente distintos, de modo que cada uno de ellos sea una potencia de cierto número primo. A cada conjunto de éstos ponemos en correspondencia una suma directa de grupos cíclicos, cuyos órdenes sean iguales a los números de este conjunto. Todos los grupos abelianos finitos obtenidos de este modo, resultan ser no isomorfos dos a dos, y cualquier otro grupo abeliano finito es isomorfo a uno de estos grupos.

INDICE ALFABETICO

KISELEV A., KRASNOV M., MAKARENKO G.
PROBLEMAS DE ECUACIONES DIFERENCIALES ORDINARIAS

Los autores de este libro son Mijail Krasnov, *Grigori Makarenko, candidatos a doctores en ciencias* físico-matemáticas y docentes del Instituto Energético de Moscú, y Alejandro Kiselev, colaborador científico superior del Instituto Unificado de investigaciones nucleares de la ciudad de Dubno.

En este libro se han recopilado cerca de 1 000 problemas y ejercicios del curso de ecuaciones diferenciales ordinarias.

Se ha incluido también el método de isoclinas para las ecuaciones de I y II orden, problemas para hallar las trayectorias ortogonales, dependencia e independencia lineales de los sistemas de funciones. Además, contiene problemas para hallar la estabilidad de las soluciones, el método del parámetro pequeño, el método para resolver ecuaciones y los sistemas.

Cada parágrafo empieza con una breve introducción teórica. Después se exponen las determinaciones y métodos principales de solución de los problemas. Todos los problemas van acompañados de su resultado; para algunos de ellos hay indicaciones de cómo resolverlos.

En esta obra se ha incluido también cierta cantidad de problemas muy complejos.

Es un libro de texto para los estudiantes de los centros de enseñanza superior. Ha aparecido dos veces editado en ruso.

Formato 13,5 × 20,5 cm. Encuadernado en tela. 208 págs.